



UNIVERSIDADE DE ÉVORA
ESCOLA DE CIÊNCIAS SOCIAIS

Mestrado em Economia e Gestão Aplicadas
Especialização Agro-Negócio

Dissertação

**Um modelo de desagregação de encargos de produção variáveis
da Base RICA por actividade**

Maria do Socorro Soares Rosário

Orientador:

Maria Leonor Pimenta Marques Verdete da Silva Carvalho

Co-Orientador:

Rui Manuel de Sousa Fragoso

Setembro de 2011

Mestrado em Economia e Gestão Aplicadas
Especialização Agro-Negócio

Dissertação

**Um modelo de desagregação de encargos de produção
variáveis
da Base RICA por actividade**

Maria do Socorro Soares Rosário

Orientador:

Maria Leonor Pimenta Marques Verdete da Silva Carvalho

Co-Orientador:

Rui Manuel de Sousa Fragoso

*Eles não sabem, nem sonham,
que o sonho comanda a vida.
Que sempre que um homem sonha
o mundo pula e avança
como bola colorida
entre as mãos de uma criança.*

António Gedeão

Quando se tem um objectivo e se há determinação,
o sonho transforma-se, a ideia desenvolve-se e a obra surge.

Agradecimentos

Aos meus orientadores,

Prof. Doutora Maria Leonor da Silva Carvalho e Prof. Doutor Rui Fragoso,

pela disponibilidade, pelo apoio, pelas sugestões e pelo carinho que sempre manifestaram.

Um modelo de desagregação de encargos de produção variáveis da base RICA por actividade

Resumo

A informação desagregada relativamente a factores de produção é um problema generalizado das ciências agrárias. Recorre-se à utilização de inquéritos directos e pessoais junto do agricultor, com recurso a amostras pequenas, que são dispendiosos e demorados. O uso de técnicas e métodos alternativos pode ser uma experiência válida para estimar os custos variáveis de produção com um menor encargo. A base de dados da Rede de Informação de Contabilidades Agrícolas (RICA) é uma fonte de informação útil para estudar aspectos da política agrícola, respeita o total dos custos por tipo de factor não sendo desagregada por actividade. A obtenção dos custos variáveis por actividade é extremamente importante para a gestão da exploração e para a análise da política agrícola.

Esta dissertação tem como principal objectivo estimar os coeficientes de afectação dos custos às actividades agrícolas, a partir da RICA: é desenvolvido um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia, aplicado na região do Alentejo e usando os dados da base RICA de 2008.

Os resultados mostram que a utilização da teoria da entropia é uma opção adequada para a estimação de coeficientes de afectação dos custos dos factores de produção às actividades num contexto de informação incompleta. Os estimadores de entropia apresentam de um modo geral bons resultados do ponto de vista estatístico e econométrico, mas os modelos de Mínima Entropia Cruzada permitem obter resultados mais aderentes à realidade do que os modelos de Máxima Entropia Generalizada.

Palavras-chave: máxima entropia; custos variáveis; modelo de desagregação; base RICA

A model of disaggregation of costs of production variables from FADN data base per activity

Abstract

The lack of disaggregated data on factors of production is a widespread problem in the agricultural sciences. Usually, to obtain such information we do direct and personal surveys, using small samples. As these surveys are expensive and time consuming, the use of alternative techniques and methods to estimate the variable costs of production, with significantly lower costs, may be a worthwhile experience.

The Farm Accounting Data Network (FADN) is a very useful source of data for studying agricultural policy. However FADN data concerns total costs by type of cost, being not broken down by activity. The attainment of variable unit costs of production by activity is extremely important, not only from the standpoint of business management, but also in the context of agricultural policy analysis.

To mitigate the lack of information on input costs in agriculture, the main goal of this paper is to estimate good coefficients for allocating those costs to farming activities, from the FADN database. We develop a model based on the theory of maximum entropy to breakdown the input variable costs by activity. This model is applied in the context of the Alentejo agricultural region, using FADN 2008 data.

The results show that the use of theory of entropy is an appropriate method for the estimating of coefficients for allocation costs of factors of production activities in a context of incomplete information. The entropy estimators have generally good results in terms of statistical and models but the Minimum Cross Entropy model results allow more adherent to reality than the Generalized Maximum Entropy models.

Keywords: maximum entropy, variable costs, FADN database, disaggregation model

Índice

Agradecimentos	vii
Resumo	ix
Abstract	x
Índice de Tabelas	xiii
Índice de Figuras	xv
Acrónimos	xvii
1. Introdução	1
1.1 - Objectivos	3
1.2 - Organização do Trabalho	4
2. Desagregação de Custos de Produção e Estimação de Coeficientes	7
2.1 - A formalização do problema de desagregação dos custos da produção agrícola	7
2.2 - Métodos utilizados	8
2.3 - Método de Máxima Entropia	10
2.3.1 - A Máxima Entropia Generalizada	12
2.3.2 - A Mínima Entropia Cruzada	14
3. Metodologia	17
3.1 - O Modelo Tobit de Máxima Entropia Generalizada	17
3.2 - O Modelo de Mínima Entropia Cruzada	20
3.3 - Limites para os intervalos dos vectores de suporte	21
3.4 - Indicadores de qualidade e precisão	22
4. Informação Disponível e Implementação Empírica do Modelo	25
4.1 - Caracterização da RICA	25
4.2 - Estratificação do campo de observação	30
4.2.1 - Ponderação	32
4.2.2 - Comparabilidade e Controlo	32
4.3 - A RICA em Portugal	34
4.3.1 - Os encargos de exploração	36
4.3.2 - Instrumentos de trabalho	38
4.4 - Organização dos dados e construção do modelo	39
4.4.1 - Selecção das explorações, actividades e itens de custo	40
4.4.2 - Organização dos dados	44
4.4.3 - Caracterização das empresas seleccionadas	44

5. Resultados	51
5.1. - Indicadores de Entropia Normalizada e Capacidade de Prescrição Pseudo- R^2	52
5.2 - Ganhos de informação desagregada.....	55
5.3 - Validação das Estimativas dos Coeficientes de Afecção dos Custos.....	57
5.4 - Discussão de resultados	63
6. Conclusões e recomendações	65
Bibliografia.....	69
ANEXOS.....	75
Anexo 5.1.- Exemplo de uma parte de listagem do GAMS	77
Anexo 5.2.- Análise descritiva dos desvios absolutos médios	81

Índice de Tabelas

Tabela 4.1- Amostra RICA 2008: n.º de explorações agrícolas por classes de OTE e DE	35
Tabela 4.2- Amostra RICA 2008: n.º de explorações agrícolas por classes de DE	36
Tabela 4.3- N.º de explorações por classes de OTE e MBT no Alentejo.....	41
Tabela 4.4 - Caracterização das explorações da OTE 1	45
Tabela 4.5 - Caracterização das explorações da OTE 3	47
Tabela 4.6 - Caracterização das explorações da OTE 4 componente vegetal.....	48
Tabela 4.7 – Caracterizaçã3o das explorações da OTE 4 componente animal	49
Tabela 5.1.- Resultados de Entropia normalizada e pseudo-R ² para os modelos de MEG e ECG das explorações agrícolas da OTE culturas arvenses	52
Tabela 5.2.- Resultados de Entropia normalizada e pseudo-R ² para os modelos de MEG e ECG das explorações agrícolas da OTE culturas permanentes.....	53
Tabela 5.3.- Resultados de Entropia normalizada e pseudo-R ² para os modelos de MEG e ECG das explorações agrícolas da OTE herbívoros-componente vegetal	54
Tabela 5.4.- Resultados de Entropia normalizada e pseudo-R ² para os modelos de MEG e ECG das explorações agrícolas da OTE herbívoros-componente animal.....	55
Tabela 5.5.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE culturas arvenses.....	55
Tabela 5.6.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE culturas permanentes	56
Tabela 5.7.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente vegetal	57
Tabela 5.8.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente animal.....	57
Tabela 5.9.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas arvenses	58
Tabela 5.10.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas arvenses.....	59
Tabela 5.11.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas permanentes	60
Tabela 5.12.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas permanentes	60
Tabela 5.13.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente vegetal	61

Tabela 5.14.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente vegetal	62
Tabela 5.15.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente animal	62
Tabela 5.16.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente animal	63
Tabela A 5.1.- Análise descritiva do DAM em MEG nas Culturas Arvenses	81
Tabela A 5.2.- Análise descritiva do DAM em ECG nas Culturas Arvenses	81
Tabela A 5.3.- Análise descritiva do DAM em MEG nas Culturas Permanentes.....	81
Tabela A 5.4.- Análise descritiva do DAM em ECG nas Culturas Permanentes	81
Tabela A 5.5.- Análise descritiva de DAM em MEG nos Herbívoros componente Vegetal.....	82
Tabela A 5.6.- Análise descritiva de DAM em ECG nos Herbívoros componente Vegetal	82
Tabela A 5.7.- Análise descritiva de DAM em MEG nos Herbívoros componente Animal	82
Tabela A 5.8.- Análise descritiva de DAM em ECG nos Herbívoros componente Animal.....	82

Índice de Figuras

Figura 4.1- Amostra RICA estratificada por classes de OTE, DE e Região	331
Figura 4.2- Representação do universo RICA	33

Acrónimos

CE- Comunidade Europeia

CEP- Custos específicos proporcionais

CN- Cabeça normal

DAM- Desvio absoluto médio

DE- Dimensão económica

DGAgri- Direcção Geral de Agricultura e Desenvolvimento Rural da EU em Bruxelas

DIG- Disaggregation Informational Gain

EA- Explorações Agrícolas

CE- Comunidade Europeia

CEP- Custos específicos proporcionais

CN- Cabeça normal

DAM- Desvio absoluto médio

DE- Dimensão económica

DGAgri- Direcção Geral de Agricultura e Desenvolvimento Rural da EU em Bruxelas

DIG- Disaggregation Informational Gain

EA- Explorações Agrícolas

EC- Mínima Entropia Cruzada

ECG- Entropia Cruzada Generalizada

EU- União Europeia

FADN- Farm Accounting Data Network

FResult- Resultados Padrão

GAMS- General Algebraic Modeling System

GID- Ganhos de Informação Desagregada

GPP- Gabinete de Planeamento e Políticas

IE 2005- Inquérito à Estrutura das Explorações Agrícolas de 2005

MAMAOT- Ministério de Agricultura, Mar, Ambiente e Ordenamento do Território

MBP- Margem Bruta Padrão

ME- Metodologia da Máxima Entropia

MEG- Máxima Entropia Generalizada

OTE - Orientação Técnica Económica

PAC- Política Agrícola Comum

PB- Produção Bruta

RICA- Rede de Informação de Contabilidades Agrícolas

SAU- Superfície Agrícola Utilizada

Tipologia- Tipologia para a Classificação Comunitária das Explorações Agrícolas

UDE- Unidades de Dimensão Europeia

VPP- Valor de Produção Padrão

1. INTRODUÇÃO

Na União Europeia há uma grande diversidade de estruturas e sistemas de produção e, apesar da crescente tendência para a especialização, a maioria das explorações agrícolas são multiproduto.

Para facilitar a análise e para haver comparabilidade das características estruturais e dos resultados económicos das explorações agrícolas dos diferentes Estados Membros, desenvolveu-se uma Rede de Informação de Contabilidades Agrícolas (RICA) com o objectivo de recolher anualmente toda a informação relativa à contabilidade da exploração agrícola, e a Tipologia para a Classificação Comunitária das Explorações Agrícolas (Tipologia) pertencentes à RICA.

Em Portugal, desde os anos 80 que a RICA¹ tem desenvolvido um trabalho de recolha de informação a nível da exploração agrícola, numa amostra constante conforme o Regulamento n.º 79/65/CEE que institui a RICA e o Regulamento n.º 2143/81/CEE que a melhora. A informação disponível baseia-se nos dados de cada exploração agrícola numa visão contabilística, em que os custos totais por tipo de factor são agregados, não sendo detalhados por actividade.

Aquando do desenvolvimento da base de dados RICA, estava prevista a recolha de informação de modo a permitir a obtenção não só dos resultados contabilísticos por exploração, mas também de dados sectoriais por actividade e por tipo de encargo, em todas as regiões e explorações. Deste modo, de uma forma rápida e segura, poder-se-ia chegar a todos os detalhes das actividades da exploração e melhor controlar a sua gestão. Com o decorrer do tempo e devido às dificuldades sentidas, quer pelo agricultor, quer pelo técnico que o acompanha na recolha de informação, relativamente à atribuição de valores aos encargos por actividade, essa informação deixou de ser detalhada e por conseguinte ficou gorada a expectativa da sua utilização.

A recolha de informação, tratamento e divulgação do trabalho designado como “Contas de cultura” ou melhor custos de produção, estava a ser feita com o detalhe de todos os factores de produção (quantidade e custos) que intervêm na produção de uma determinada actividade. Devido aos custos e à exigência dos dados envolvidos e aos novos interesses das instituições a que os técnicos

¹ Serviço pertencente ao Ministério de Agricultura, Mar, Ambiente e Ordenamento do Território (MAMAOT)

pertenciam, a publicação “Contas de cultura” foi gradualmente ficando para segundo plano, até que se deixou de fazer. Posteriormente também a recolha de informação se deixou de realizar.

A Tipologia Comunitária, estabelecida pela Decisão 78/463/CEE para classificar as explorações da RICA, requeria a elaboração da Margem Bruta Padrão (MBP) por actividade, em que ao Produto Bruto eram retirados os Custos Específicos Proporcionais. Esta metodologia foi alterada, passando a ser exigida apenas a determinação da Produção Bruta, ficando desconhecidos os encargos específicos por actividade.

Sendo a competitividade um objectivo amplamente divulgado e solicitado às empresas do sector agrícola, ela não pode ser alcançada apenas por via dos preços de comercialização de produto, pois as empresas compradoras querem cada vez mais pagar menos, independentemente da origem do produto. A forma de aumentar a competitividade nesse contexto consiste em melhorar a eficiência (gestão) na utilização dos factores de produção nas diversas actividades.

A análise económica das actividades agrícolas em resposta às mudanças de medidas de políticas e de mercado, sistemas de produção, tecnologias e competitividade também se baseiam na estrutura de custos de produção das actividades. Estas estruturas são inerentes à existência de uma empresa agrícola e esses custos podem ser variáveis, fixos e atribuídos, sendo os primeiros os custos específicos de uma actividade. A informação relativa aos encargos fixos e variáveis de produção por unidade de actividade é extremamente importante, não só ao nível da gestão da empresa agrícola mas também na análise de políticas.

Para produzir e actualizar essa informação de base é necessário muito capital, quer humano quer financeiro, que actualmente não é possível dispor. A recolha de dados ao nível da exploração é difícil, exige deslocações frequentes e inquéritos feitos com muita dedicação e tempo, quer por parte dos técnicos, quer por parte dos agricultores.

O agricultor, apesar de dispor da informação requerida, não consegue detalhá-la na forma contabilística mais amigável para a análise económica da empresa e a gestão das suas actividades, e

que permita saber quais as actividades mais interessantes a desenvolver e as que podem ser alteradas.

As explorações agrícolas (EA) que fazem parte da RICA têm essa informação agregada para a exploração por tipo de factores de produção e nunca separados por actividade, ou seja, as actividades têm a sua produção individualizada mas os custos dos factores de produção estão organizados de uma forma agregada para a totalidade da exploração. Desta forma não é possível conhecer a estrutura de custos e, por conseguinte, a rentabilidade por actividade.

No momento em que as flutuações nos preços de factores são enormes, essa falta de informação atinge uma dimensão muito grande e surge a necessidade urgente de a colmatar. A existência de informação recolhida na empresa com custos agregados por factor e receitas/produções detalhadas por actividade incita a utilizar essa informação para a sua desagregação.

O uso de técnicas alternativas baseadas em métodos econométricos ou na metodologia da máxima entropia pode ser uma solução válida para estimar os encargos com um custo muito baixo.

Esta necessidade já foi sentida por diversos autores, nomeadamente, Lence & Miller (1998) que propuseram um modelo de desagregação de custos de factores e produções baseado na metodologia da Máxima Entropia (ME). Também Peeters & Surry (2002; 2005) no caso das explorações de Saskatchewan no Canadá e no caso nas explorações agrícolas em Brittany, França recorreram a uma abordagem baseada na entropia para proceder à estimação do custo dos factores.

1.1 - Objectivos

Para tentar responder ao problema da falta de informação relativa aos custos dos factores por actividade agrícola, esta dissertação tem como principal objectivo estimar os coeficientes de afectação desses custos às actividades agrícolas, a partir da base de dados da RICA.

Para o efeito é desenvolvido um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia. Este modelo é aplicado no

contexto da região agrária do Alentejo, segundo os dados da RICA de 2008. Desta forma, será também fornecida informação sobre custos variáveis por actividade, por exploração e a nível regional.

Para além dos aspectos metodológicos relacionados com o desenvolvimento de uma ferramenta eficaz de desagregação dos custos dos factores por actividade agrícola, esta dissertação encerra como objectivos específicos a avaliação das principais propriedades estatísticas dos estimadores da entropia e respectiva validação pela comparação dos seus resultados com os dados observados.

Este estudo pretende contribuir para mitigar a falta de informação sobre encargos variáveis por actividade, necessária para a análise de políticas agrícolas e para a elaboração e monitorização de novas medidas de política agrícola e de desenvolvimento rural, portanto, a sua importância para a RICA é de algum modo significativa devido a:

- a) Permitir a desagregação dos encargos de produção variáveis da base RICA por actividade, por exploração e por região;
- b) Economizar esforços humanos e financeiros com a recolha de informação junto aos agricultores;
- c) Fazer a ponte entre o trabalho que já é desenvolvido e novas acções a serem desencadeadas com a base RICA.

1.2 - Organização do Trabalho

Tendo em conta os objectivos propostos, a dissertação encontra-se organizada em seis capítulos.

O Capítulo 1 corresponde a uma introdução ao tema com as necessidades sentidas, os fins para os quais essas necessidades são reconhecidas como essenciais, os objectivos a alcançar com a dissertação e de que modo é que os fins propostos podem colmatar as necessidades e melhorar o panorama da informação.

O Capítulo 2 é um capítulo de revisão bibliográfica que tem como principais objectivos proceder à formalização do problema da desagregação dos custos de produção agrícola, apresentar os principais métodos utilizados para esse efeito e ideias do referencial quadro teórico do método da máxima entropia.

O Capítulo 3 diz respeito à concepção e desenvolvimento de um instrumento de análise baseado na máxima entropia, nomeadamente, um modelo de máxima entropia generalizada, um modelo de entropia cruzada generalizada e a definição critérios de avaliação das propriedades estatísticas dos estimados e da sua validação.

O Capítulo 4 destina-se à implementação empírica do modelo tendo por base a RICA que fornece a informação de base. Neste capítulo procede-se também à organização dos dados para serem utilizados na aplicação no modelo.

O Capítulo 5 é dedicado à apresentação e discussão de resultados. Nele são apresentados os resultados dos testes estatísticos que foram realizados aos estimadores da entropia e são comparados os valores dos estimadores com os dados reais disponíveis.

Por último, no Capítulo 6, são apresentadas as principais conclusões e recomendações, são referidas as principais limitações metodológicas sentidas no decorrer do estudo e são identificados desenvolvimentos futuros.

2. DESAGREGAÇÃO DE CUSTOS DE PRODUÇÃO E ESTIMAÇÃO DE COEFICIENTES

Neste capítulo procede-se a uma breve revisão dos principais métodos utilizados na desagregação de custos, dando especial atenção aos métodos baseados na teoria da Máxima Entropia. Começa-se por descrever o problema da desagregação dos custos de produção em termos da sua formulação matemática. Em seguida referem-se os principais métodos utilizados para tratar esse tipo de problema, dedicando-se a última parte do capítulo aos métodos da entropia, nomeadamente, a formulação da Máxima Entropia Generalizada e a formulação da Mínima Entropia Cruzada.

2.1 - A formalização do problema de desagregação dos custos da produção agrícola

A falta de informação sobre os sistemas de produção, mesmo considerando a informação existente na base RICA, que é a fonte oficial e privilegiada nos meios da União Europeia, é um problema determinante que pode ser atenuado com o recurso a métodos de desagregação de dados, nomeadamente dos custos da produção agrícola por actividades. Deste modo o problema da falta de informação transforma-se num problema de desagregação de custos, formalizado para uma amostra de T explorações, que realizam K actividades agrícolas utilizando factores de produção que dão origem a I itens de custo:

$$X_i^t = \sum_{k=1}^K \alpha_{ik} Y_k^t + u_i^t \quad i=1,2,...,I \text{ e } t=1,2,...,T \quad (1)$$

X_i^t = Custo do *input* i por exploração t;

α_{ik} = Coeficiente de afectação de custos por *input* i e por produto k;

Y_k^t = Receita total ou produto bruto por produto k e por exploração t;

u_i^t = Resíduo aleatório da distribuição que é específico de cada *input* i e cada exploração t.

Os coeficientes de afectação α_{ik} são os parâmetros desconhecidos da equação (1) e são definidos em termos dos seus valores por actividade agrícola ou produto e por item de custo ou *input* para a globalidade das empresas da amostra objecto de estudo.

2.2 - Métodos utilizados

Existem vários métodos para estimar os custos específicos das actividades das explorações agrícolas. Estas estimativas são baseadas nas relações que se estabelecem entre os coeficientes de afectação da produção e os respectivos factores, assumindo a hipótese de que os agentes económicos assumem um comportamento de maximização do lucro e tendo em conta as devidas desagregações dos factores, estimativas de custos de produção e coeficientes de produtividade (Just *et al.*, 1983; Lence & Miller, 1998).

Para se proceder à desagregação dos custos por actividade agrícola através da estimativa de parâmetros, pode recorrer-se às técnicas de regressão linear. No entanto, a fiabilidade dos parâmetros estimados depende do número de empresas ou de explorações agrícolas da amostra. Segundo Love (1999) e Zhang & Fan (2001), quando a informação provém de uma amostra pequena, as técnicas de regressão linear não podem ser utilizadas, pelo que deverão ser adoptados outros procedimentos metodológicos.

Por outro lado, Peeters & Surry (2002; 2005) referem que a imposição de restrições contabilísticas, como a que garante a igualdade do rendimento bruto total aos custos totais, dificultam ou inviabilizam mesmo a utilização de algumas técnicas de análise para estimar os coeficientes de desagregação dos custos. Estes autores também se referem às técnicas de regressão linear e à técnica de estimação bayesiana, como as metodologias clássicas mais usadas, apesar dos problemas que ambas comportam na sua aplicação.

Nas técnicas de regressão linear, os principais problemas prendem-se com a possibilidade de obtenção de estimativas negativas e com o facto de todas as equações serem interdependentes (Bewley, 1986). O problema da negatividade dos coeficientes pode ser resolvido com a utilização

de técnicas de estimação bayesiana ou através de método dos mínimos quadrados com restrições (Moxey e Tiffin, 1994). No entanto, estes métodos, para além de não permitirem incorporar as restrições de saldo contabilístico, a sua aplicação empírica é muito complexa.

No seu estudo, Pires *et al.* (2010) concluiu que o método dos mínimos quadrados não era adequado para estimar parâmetros a partir de amostras pequenas, contrariamente ao método de máxima entropia (ME) que nessas condições permitiu melhores estimativas na desagregação de dados. Estas conclusões estão de acordo com os resultados apresentados por outros autores, como Golan & Judge (1992), Golan & Perloff (2002) e Kitamura & Stutzer (2002), que foram testemunhando ao longo dos anos a capacidade da ME para estimar a desagregação de dados económicos.

Golan *et al.* (1994) usou a ME e a mínima entropia cruzada (EC) para derivar uma matriz de contabilidade social (*social accounting matrix*). Mais tarde, Golan *et al.* (1997) e Golan *et al.* (2001) utilizaram o método da máxima entropia generalizada (MEG) para estimar os parâmetros de um modelo de regressão. Paris & Howitt (1998) recorreram a uma formulação de MEG para estimar os parâmetros das funções custo das explorações a serem incorporadas na função objectivo de um modelo de programação matemática positiva.

Leon *et al.* (1999) e Peeters & Surry (2002) mostraram que a qualidade de precisão dos coeficientes extraídos de amostras de explorações agrícolas com dados incompletos através dos métodos de entropia é elevada. Estes autores, para além de concluírem que as características estatísticas demonstradas pelos estimadores da entropia são boas, verificaram também que os resultados não apresentam diferenças significativas, quando comparados com uma amostra de encargos observados desagregados por actividade.

Fraser (2000) utilizou a MEG para estimar a procura de carne no Reino Unido e Zhang & Fan (2001) recorreram à MEG para estimar as tecnologias de produção na agricultura chinesa e para proceder à desagregação dos factores de produção por província.

Garvey & Britz (2002) estimaram a desagregação de custos a partir da contabilidade da RICA também com a MEG e Campbell & Hill (2005 e 2006) mostraram como colocar restrições de informação num modelo de MEG utilizando diferentes vectores de suporte.

Howitt & Msangi (2006) recorreram à MEG para estimar funções de produção numa amostra pequena de 27 explorações no Rio Bravo, México. Rezek & Campbell (2006) apoiaram-se na MEG para estimar os custos de diversos efluentes / poluentes a partir de dados de painel.

Hansen & Surry (2007) derivaram a procura de quantidades de factores para diferentes produções na Alemanha utilizando o método de máxima entropia aplicado aos dados das Contas Económicas Regionais

Fragoso *et al.* (2008) utilizaram um modelo EC para proceder à desagregação das superfícies agro-florestais de acordo com zonas agro-ecológicas pré-definidas para o ecossistema montado no Alentejo.

Lips em 2009, partindo da informação da RICA, utilizou a ME para desagregar os custos das várias actividades pelas diversas áreas de culturas arvenses na Suíça.

Xavier *et al.* (2010) utilizou a ME para criar um modelo de desagregação dos dados da utilização da terra, dos efectivos pecuários e de produções. Papalia (2010) também usou a MEG para a estimação de um modelo de interacção espacial a partir de dados de painel.

2.3 - Método de Máxima Entropia

A entropia de um sistema é uma medida da desordem do sistema. Este conceito foi introduzido por Shannon (1948) como uma medida de incerteza no contexto da teoria de informação, em que a informação contida num evento aleatório tem uma relação inversa com a sua probabilidade de ocorrência. Deste modo, a entropia pode ser medida em termos de probabilidades não observadas ou percebidas. Considerando um evento com $k = \{1, 2, \dots, k, \dots, n\}$ estados de decisão e com

probabilidade de ocorrência p_k , o valor de entropia H que mede a incerteza associada à distribuição das probabilidades de ocorrência do evento é dada por:

$$H(p) = - \sum_{k=1}^n p_k \ln p_k \quad (2)$$

A entropia da distribuição de probabilidade $H(p)$ é sempre positiva e cresce com a incerteza, atingindo o seu valor máximo quando $p_1 = p_2 = \dots = p_n = 1/n$, i.e., quando a distribuição de probabilidade p_k é uniforme. Assim, maximizar a entropia $H(p)$ passa por escolher o vector p_k , que melhor se ajusta à distribuição uniforme, de acordo com os dados existentes, nomeadamente, as restrições de informação. Se não existir informação sobre uma distribuição de probabilidade particular, então deveria ser considerado que todos os resultados podem ser igualmente prováveis e isso corresponde à maximização da função entropia (Pires *et al.*, 2010).

O princípio de ME proposto por Jaynes (1957) generaliza a ideia de que existe uma distribuição de probabilidade p_k , que num determinado momento e em determinadas condições, pode maximizar a entropia. Este critério é racional e lógico para escolher uma função de distribuição de probabilidades específica que maximize a entropia, utilizando uma determinada restrição de informação. Deste modo, p_k é a probabilidade de ocorrência do evento k em que a medida de entropia está limitada pela informação existente. De acordo com Soofi (2000) e Golan e Perloff (2002), este princípio só usa a informação considerada relevante. O princípio da ME pode ser apresentado através da seguinte formulação:

$$\text{Max } H(p) = - \sum_{k=1}^n p_k \ln p_k \quad (3)$$

Sujeito a

$$\sum_{k=1}^n p_k y_k = X \quad (4)$$

$$\sum_{k=1}^n p_k = 1 \quad (5)$$

$$p_k \geq 0 \quad (6)$$

Onde, p_k que representa a probabilidade de ocorrência do evento é a variável desconhecida a determinar, Y_k é uma variável independente e X é uma variável dependente. Neste caso, os valores de Y_k e de X são conhecidos e a sua relação em função de p_k , e por conseguinte, o valor máximo da função H em (3) estão condicionados pelas restrições de informação que se estabelecem em (4). As expressões (5) e (6) representam as propriedades das probabilidades e a condição de não negatividade, respectivamente. A primeira garante que a soma das probabilidades estimadas é igual à unidade e a segunda que os valores das probabilidades estimadas são positivos ou nulos.

De acordo com o princípio da ME, há apenas uma distribuição de probabilidade que pode maximizar a medida de entropia, sendo consistente com a informação disponível. Essa distribuição é aquela que compreende a condição de consistência de informação com um mínimo de conteúdo.

A partir da informação disponível, a ME vai proceder à procura da distribuição de probabilidade, i.e., a oportunidade que uma variável pode ter ao longo de um espaço, sujeita às restrições impostas e que maximize a informação. Essa distribuição é uma função em que os valores da variável fazem parte do domínio e a imagem é o conjunto de probabilidades que essa variável pode assumir em cada ponto do domínio. O conjunto da imagem está restrito ao intervalo entre zero e um (probabilidade). Segundo Jaynes (1968) e Silva (2008), o viés mínimo da distribuição de p_k é o que maximiza a entropia sujeita a determinada informação ou que satisfaça as condições do princípio de máxima entropia.

2.3.1 - A Máxima Entropia Generalizada

Apesar de todas as virtudes do critério da ME, Golan e Judge (1992) propuseram e utilizaram uma generalização que corresponde ao método da Máxima Entropia Generalizada. Esta formulação permite o tratamento da componente do ruído e os coeficientes não observados são relacionados como um conjunto de vectores de suporte conhecidos e discretos.

Na formulação da MEG, a expressão referida em (1) que traduz o problema de desagregação de custos, tem de ser reparametrizada para acomodar a formulação do erro da distribuição da variável

X e os vectores de suporte que introduzem informação prévia na estimativa. Deste modo, da reparametrização de (1) obtêm-se os seguintes vectores α e u :

$$\alpha_{ik} = \sum_{m=1}^M z_{i,k}^m p_{i,k}^m \quad \forall i \in 1, 2, \dots, I ; \quad \forall k \in 1, 2, \dots, K \quad (7)$$

$$u_i^t = \sum_{n=1}^N v_{i,t}^n w_{i,t}^n, \quad \forall i \in 1, 2, \dots, I ; \quad \forall t \in 1, 2, \dots, K \quad (8)$$

Onde, $z_{i,k}^m$ é o vector de suporte de dimensão M para k produtos com encargos i associados; $v_{i,t}^n$ é o vector de suporte de dimensão N para T termos de erros associados a u com encargos i ; $p_{i,k}^m$ e $w_{i,t}^n$ são os vectores da distribuição desconhecida das probabilidades de ocorrência dos k estados de X e do erro u que lhe está associado.

A incorporação das variáveis reparametrizadas no critério da ME dá lugar ao seguinte modelo de MEG, que é agora apresentado na forma matricial:

$$\text{Max } H(p, w) = -p' \ln p - w' \ln(w) \quad (9)$$

Sujeito a

$$x = \alpha y + u = z' p y + v' w \quad (10)$$

$$(I_K \otimes i_M) p = i_K \quad (11)$$

$$(I_T \otimes i_N) w = i_T \quad (12)$$

Onde x , y , z e v são conhecidos, p e w são os vectores de probabilidade desconhecidos a serem estimados e \otimes é o produto Kronecker. A função da máxima entropia (9) é agora uma função da entropia conjunta da probabilidade dos parâmetros e do erro da distribuição, a restrição de informação (10) na relação entre a variável dependente x e independente y tem em conta a estimativa da probabilidade do erro w ao longo de um vector suporte de dimensão N e da probabilidade dos parâmetros p ao longo de um vector predefinido de dimensão M . As condições (11 e 12) limitam a soma das probabilidades p e w ao valor unitário. Nestas condições, a função de

máxima entropia (9) apresenta independência entre os vectores \mathbf{p} e \mathbf{w} quando sujeitos às restrições (10), (11) e (12).

De acordo com as condições de optimização de primeira ordem da função de Lagrange, os vectores \mathbf{p} e \mathbf{w} podem ser calculados por:

$$p_{km} = \frac{\exp z_{km} y'_k \lambda_t}{\sum_{m=1}^M \exp z_{km} y'_k \lambda_t} \quad k = 1, 2, \dots, K \quad (13)$$

$$w_{tn} = \frac{\exp v_{tn} \lambda_t}{\sum_{n=1}^N \exp v_{tn} \lambda_t} \quad t = 1, 2, \dots, T \quad (14)$$

Onde $\hat{\lambda}$ é o vector multiplicador de Lagrange para as restrições de informação, que associadas a \mathbf{p} e \mathbf{w} , satisfazem as condições (10), (11) e (12). Este vector incorpora a função de informação $F \lambda_t$, dada pela matriz Hessiana da função objectivo em (9):

$$\nabla_{\mathbf{p}, \mathbf{w}} \mathbf{p}', \mathbf{w}' F \mathbf{p}, \mathbf{w} = \begin{pmatrix} \mathbf{p}^{-1} & 0 \\ 0 & \mathbf{w}^{-1} \end{pmatrix} \quad (15)$$

onde ∇ é uma matriz definida positiva para $\mathbf{p} > 0$ e $\mathbf{w} > 0$, que satisfaz a condição de estrita convexidade.

Nestas condições pode-se concluir que o problema tem uma solução única. O facto das estimativas da função de Lagrange para \mathbf{p} e \mathbf{w} serem dadas na forma exponencial é a garantia de que $p_{i,n}$ e $w_{i,n}$ são sempre positivos.

2.3.2 - A Mínima Entropia Cruzada

A metodologia da entropia permite incorporar na estimativa informação adicional à contida na amostra da variável X através da utilização do critério da mínima entropia cruzada. Este princípio foi introduzido por Good (1963) e é muito útil para melhorar o resultado das estimativas e diminuir as discrepâncias relativamente aos dados observados.

A entropia cruzada é uma medida de afastamento entre as estimativas das probabilidades *a posteriori* p e o *prior* (antecedente) de informação q . Este vector q é o vector das probabilidades conhecidas, i.e., o vector da informação prévia que é adicionada à estimativa para além da informação que está contida na amostra da variável X . O objectivo é minimizar a distância entre as duas distribuições. As informações prévias que constituem o prior de informação são aquelas que realmente existem ou são consideradas como reais de modo a poder ser possível obter estimativas mais reais. Neste caso, o valor da entropia atinge o seu mínimo quando a distribuição do vector p desconhecido for igual à distribuição do prior de informação q .

A formulação matemática do critério da minimização da entropia cruzada pode ser dada pelas seguintes expressões:

$$\text{Min } EC(p, q) = \sum_{k=1}^K p_k \ln \frac{p_k}{q_k} = - \sum_{k=1}^K p_k \ln p_k + \sum_{k=1}^K p_k \ln q_k \quad (16)$$

Sujeito a

$$\sum_{k=1}^K p_k y_k = X \quad (17)$$

$$\sum_{k=1}^K p_k = 1 \quad (18)$$

$$p_k \geq 0 \quad (19)$$

Onde, p_k é a probabilidade a estimar e q_k é a probabilidade conhecida, i.e., o vector de informação existente.

O princípio da entropia cruzada também pode ser formulado como um problema de entropia cruzada generalizada. Golan *et al.* (1996) estendeu a formulação da MEG ao critério da mínima entropia cruzada, o que permitiu levar em consideração a estimativa da distribuição das probabilidades e a medição dos erros.

3. METODOLOGIA

Para proceder à desagregação dos custos de produção da exploração agrícola por actividade a partir da base de dados da RICA utilizou-se o método da ME. Para o efeito foi considerado um modelo de MEG, cujos resultados foram melhorados com ajuda de um prior de informação numa formulação de entropia cruzada generalizada (ECG). A aplicação desses modelos ME permite estimar os coeficientes de afectação dos custos dos factores de produção da exploração agrícola pelas suas actividades através de um sistema de equações lineares. Este conjunto de equações não é mais do que as restrições de informação que condicionam as estimativas dos coeficientes de afectação. Os custos dos factores são assumidos como uma função linear do valor da produção das actividades.

A escolha da metodologia recaiu sobre a ME porque este método permite estimar parâmetros sem a imposição de restrições à distribuição da probabilidade e dos erros a partir de amostras de tamanho reduzido (Golan *et al.*, 1996; Campbell & Hill, 2005, 2006). Este método tem também a vantagem de não requerer informação sobre o comportamento do agricultor e de poder integrar no modelo informação adicional que não consta da amostra.

Este capítulo dedicado à metodologia está organizado em quatro partes. Na primeira parte apresenta-se a estrutura do modelo Tobit de MEG. Na segunda parte procede-se à introdução no modelo de um prior de informação e por conseguinte na transformação da formulação de MEG num modelo de ECG. Na terceira parte são discutidos os limites dos intervalos que enquadram os vectores de suporte dos parâmetros e dos erros. Por último, na parte final do capítulo são apresentados os indicadores utilizados na avaliação da qualidade do modelo.

3.1 - O Modelo Tobit de Máxima Entropia Generalizada

A metodologia da máxima entropia generalizada (MEG) tem sido muito utilizada para a resolução de problemas relacionados com a desagregação de factores. Por exemplo, Peeters & Surry (2002 e 2005) adoptaram este método para estimar os coeficientes de afectação dos custos dos factores de produção às actividades agrícolas a partir de uma amostra de explorações. Golan *et al.*, (2001) e

Peeters & Surry (2002, 2005) utilizaram esta metodologia tendo introduzido uma formulação Tobit nos seus modelo de ME. Este procedimento permite separar as situações em que a informação está disponível daquelas em que não existe. A sua aplicação é muito útil no caso da aplicação dos modelos de ME ao contexto das explorações agrícolas, na medida em nem todas as explorações realizam todas as actividades agrícolas disponíveis, correspondendo, por conseguinte, nestes casos, a uma situação de ausência de informação, que é necessário corrigir. Deste modo, tendo em conta a formulação Tobit, as equações de informação na sua forma matricial e vectorial passam a ser as seguintes:

$$X = YA + u = \begin{matrix} x_1 > 0 \\ x_2 = 0 \end{matrix} = \begin{matrix} y_1 \\ y_2 \end{matrix} A + \begin{matrix} u_1 \\ u_2 \end{matrix} \quad (20)$$

Onde x_1 , y_1 e u_1 dizem respeito aos respectivos valores não nulos da informação observada e x_2 , y_2 e u_2 representam a ausência de informação.

De acordo com Golan *et al.* (1996b) e Peeters & Surry (2002), o modelo estatístico linear pode ser alterado, reparametrizando A e u em função dos seus parâmetros aleatórios p e w_1 e w_2 e dos vectores de suporte z_1 , z_2 , v_1 e v_2 :

$$\begin{matrix} x_1 > 0 \\ x_2 = 0 \end{matrix} = \begin{matrix} y_1 \\ y_2 \end{matrix} z'p + \begin{matrix} v'_1 & w_1 \\ v'_2 & w_2 \end{matrix} \quad (21)$$

A seguir apresenta-se a formulação Tobit simplificada do modelo de MEG que foi utilizado para estimar os coeficientes de afectação dos custos dos factores de produção agrícola.

$$\begin{aligned} \text{Max } p, w_1, w_2 = & \\ - \sum_{i=1}^I \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \ln p_{k,i}^m - & \\ \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w1_{t,i}^n \ln w1_{t,i}^n - \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w2_{t,i}^n \ln w2_{t,i}^n & \end{aligned} \quad (22)$$

Sujeito a:

$$X1_i^t = \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m \cdot y1_k^t + \sum_{n=1}^N w1_{i,t}^n \cdot v1_{i,t}^n, \quad \forall t1 e \forall i \in 1, \dots, T \quad (23)$$

$$X2_i^t = \sum_{k=1}^K \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m \cdot y2_k^t + \sum_{n=1}^N w2_{i,t}^n \cdot v2_{i,t}^n, \quad \forall t \in 1, \dots, T \quad (24)$$

$$\sum_{m=1}^M p_{k,i}^m = 1 \quad \forall k \in 1, 2, \dots, I \quad (25)$$

$$\sum_{n=1}^N w1_{i,t}^n = 1, \quad \forall i \in 1, 2, \dots, I \text{ e } \forall t \in 1, 2, \dots, T_1 \quad (26)$$

$$\sum_{n=1}^N w2_{i,t}^n = 1, \quad \forall i \in 1, 2, \dots, I, \forall t \in 1, 2, \dots, T_2 \text{ e } T_1 + T_2 = T \quad (27)$$

$$\sum_{i=1}^I \sum_{m=1}^M p_{k,i}^m \cdot z_{k,i}^m = 1, \quad \forall k \in 1, 2, \dots, K \quad (28)$$

A equação (22) diz respeito à maximização da entropia conjunta da distribuição de probabilidades dos coeficientes de afectação dos custos dos factores de produção $p_{k,i}^m$ e das distribuições de probabilidades dos erros $w1_{t,i}^n$ e $w2_{t,i}^n$.

As equações (23) e (24) representam as restrições de consistência da informação, referindo-se a primeira às t_1 explorações com valores positivos em $y1_k^t$ e a segunda às t_2 explorações com valores nulos em $y2_k^t$. Essas equações compreendem a necessidade de consistência de informação com um mínimo de conteúdo, tratando as relações da equação linear do problema (1) como um sistema independente de equações, em que todos os factores de produção i são contabilizados em conjunto e em simultâneo. No entanto os valores positivos são separados dos valores nulos.

As equações (25) a (27) são restrições relativas às propriedades da probabilidade e criam no modelo a normalização dos valores de probabilidade de $p_{k,i}^m$ e $w_{i,t}^n$ relativamente às suas dimensões M e N , respectivamente. Garante-se deste modo que a soma das probabilidades estimadas é igual à unidade.

A equação (28) corresponde a uma condição adicional que estabelece que o somatório dos coeficientes de afectação dos custos dos factores de produção estimados para cada actividade k seja igual a 1.

3.2 - O Modelo de Mínima Entropia Cruzada

Para tentar melhorar os resultados das estimativas dos coeficientes de afectação dos custos, utilizou-se um prior de informação exterior à amostra de explorações agrícolas objecto de estudo, recorrendo à formulação ECG. Neste caso, o prior de informação utilizado diz respeito à estrutura dos custos utilizada pelo Gabinete de Planeamento e Políticas (GPP)², para realizar o apuramento das margens brutas padrão das actividades agrícolas para o ano de 2004.

Esse procedimento tem sido muito utilizado. Harris (2002) utilizou técnicas de mínima entropia cruzada para reconciliar dados referentes a diferentes fontes na estimativa da matriz contabilística social regional no México, onde as regiões rurais são diferenciadas por sistemas de produção agrícola com diversas tecnologias. Fragoso *et al.* (2008) mostrou que a utilização da mínima entropia cruzada pode ser um processo muito útil para a desagregação espacial de dados. Xavier *et al.* (2010) utilizou a mínima entropia cruzada para criar um modelo de desagregação dos dados da agricultura da região Algarve para um nível local numa zona de intervenção florestal.

Neste caso, o modelo de ECG tem a mesma estrutura de equações do que o modelo de MEG, ou seja, inclui as restrições de coerência de informação que permitem relacionar as variáveis dependentes com as variáveis independentes do problema e as restrições das propriedades das probabilidades. A única alteração que se verifica é ao nível da função objectivo que passa a ser a minimização da entropia cruzada entre distribuição dos coeficientes a estimar e a distribuição dos coeficientes que constituem o prior de informação, como se representa a seguir:

$$\begin{aligned} \text{Min } \alpha, w_1, w_2 = & \sum_{i=1}^I \sum_{k=1}^K \alpha_k^i \ln \frac{\alpha_k^i}{q_k^i} + \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{1,t,i}^i \ln w_{1,t,i}^i + \\ & + \sum_{t=1}^T \sum_{i=1}^I \sum_{n=1}^N w_{2,t,n}^i \ln w_{2,t,n}^i \end{aligned} \quad (29)$$

Onde α é o coeficiente de afectação dos custos que se pretende estimar e q é o coeficiente conhecido do prior de informação utilizado.

² Serviço pertencente ao Ministério de Agricultura, Mar, Ambiente e Ordenamento do Território (MAMAOT)

3.3 - Limites para os intervalos dos vectores de suporte

A escolha dos vectores de suporte, nomeadamente, dos limites do intervalo a que pertencem os seus valores e número de pontos suporte desses vectores é sempre uma tarefa difícil e delicada, sendo necessário ter em conta não só a amplitude do intervalo, como também a simetria da distribuição e por conseguinte da forma como se distribuem os seus valores no espaço.

Golan *et al.* (1996a) considera que se aumentar a largura do intervalo dos limites do vector suporte e o número de pontos suporte, pode-se ampliar a informação produzida dos dados estimados. No entanto, Paris (1998) defende que, apesar dos valores dos vectores suporte z poderem modificar as estimativas de uma forma significativa, não há uma regra para proceder à escolha de um vector certo para a distribuição.

No que se refere ao vector de suporte do erro, é muito comum a utilização da regra dos três sigmas (3σ), que consiste em considerar o erro centrado na origem e admitir uma amplitude de 3 desvios padrão à direita e de 3 desvios padrão à esquerda (Pukelsheim, 1994). Exemplos da utilização deste procedimento para tratar o erro da distribuição são os trabalhos de Golan *et al.* (1996a), Peeters & Surry (2002), Chakir (2009), Martins *et al.* (2011), Pires *et al.* (2010) e Fragoso *et al.* (2011).

Peeters & Surry (2002 e 2005) utilizaram, respectivamente, em estudos feitos para o Canadá e para França, um vector suporte z definido no intervalo $[0,1]$, que nesses casos eram os únicos valores teoricamente aceitáveis, onze pontos de suporte ($M=11$) e a regra dos três sigmas na definição do erro com um vector de suporte v com três pontos de apoio ($N=3$).

Fragoso *et al.* (2011) definiram os limites do vector suporte z no intervalo $[0,1]$ e consideraram apenas 3 pontos de apoio ($M=3$). Para o erro consideraram a regra dos três sigmas na amplitude do intervalo do vector v . Fragoso *et al.* (2008) e Van Delden *et al.*, (2006) também consideraram o mesmo procedimento, que está de acordo com (Golan *et al.*, 1996). O processo de definição do erro seguido por esses autores também foi verificado por Chakir (2009) e Martins *et al.* (2011).

Neste estudo vai-se considerar para o vector de suporte das probabilidades estimadas no intervalo $[0,1]$ com $M=3$ pontos e $Z= \{0,0.5,1\}$ porque os resultados não pareceram ser sensíveis a outras alternativas de z . No que diz respeito à definição do erro adoptou-se a regra dos três sigmas e estabeleceram-se três pontos de apoio ($N=3$) para o vector v , i.e., $v= \{-\sigma, 0, \sigma\}$.

3.4 - Indicadores de qualidade e precisão

Para avaliar a qualidade e a precisão das estimativas dos modelos de entropia apresentados na secção anterior, considerou-se o cálculo dos indicadores da entropia normalizada, do pseudo- R^2 , dos ganhos de informação obtidos no processo de desagregação de custos e do desvio absoluto médio. Esses indicadores têm sido utilizados por vários autores para aferir as propriedades estatísticas dos estimadores da entropia e para concluir da sua validade prática, por comparação com dados observados (Peeters & Surry, 2002; Howitt & Msangi, 2006; e Fragoso & Carvalho, 2011).

O indicador de entropia normalizada (S) é definido por Golan *et al.* (1996) como a proporção de incerteza que permanece na estimativa. Este indicador é muito útil para avaliar sobre o desempenho relativo das estimativas dos coeficientes da entropia no que diz respeito à geração de nova informação e à participação da informação inicial contida nos dados nessas estimativas.

No caso deste estudo, com base em Golan *et al.* (1996b) e em Peeters e Surry (2003), considerou-se o cálculo da entropia normalizada dos coeficientes de afectação dos custos dos factores de produção S_p e o cálculo da entropia normalizada para a estimativa dos erros S_w :

$$S_p = \frac{-p' \ln p}{IK \ln M} \quad (30)$$

$$S_w = \frac{-w' \ln w}{IT \ln(N)} \quad (31)$$

Onde, I é o número de factores de produção ou itens de custo, K é número de actividades agrícolas, T é o número de explorações e M e N são o número de pontos de apoio considerando para os vectores suporte z e v .

O valor do indicador da entropia normalizada varia entre 0 e 1. O valor zero indica que não há incerteza na estimativa e que a informação dos dados está totalmente reflectida nas estimativas, não se tendo por conseguinte produzido qualquer tipo de informação nova no processo de desagregação dos dados. Pelo contrário, se o valor indicado é igual à unidade, a incerteza é máxima e a participação dos dados na informação contida nas estimativas é mínima, indicando que a produção de informação nova é máxima.

Para avaliar a capacidade de estimação do modelo utilizou-se como indicador o pseudo- R^2 que foi também utilizado com o mesmo fim por Peeters & Surry (2002), Fragoso & Carvalho (2011), Fragoso *et al.* (2008), Chakir (2009), Martins *et al.* (2011) e Xavier *et al.* (2010).

O pseudo- R^2 é definido como o quadrado do coeficiente da correlação entre os valores estimados e os valores observados:

$$R_i^2 \equiv \frac{\sum_{t=1}^T x_t^i x_t^i}{\sum_{t=1}^T x_t^i x_t^i}^2, \text{ com } x_t^i = \sum_{k=1}^K \alpha_k^i y_{k,t} \quad (32)$$

Os valores do indicador pseudo- R^2 variam entre 0 e 1, significando no primeiro caso uma reduzida capacidade de estimação do modelo e no último a capacidade estimação é máxima.

Uma das medidas da qualidade da aderência das estimativas com a realidade e simultaneamente da quantidade de informação produzida no processo de desagregação é o indicador de ganhos de informação desagregada (GID). Este indicador foi estabelecido por Howitt & Reynaud (2002) sob a designação original de Disaggregation Informational Gain (DIG) e entre nós foi utilizado por Xavier *et al.*, (2010) e por Fragoso *et al.*, (2008). O cálculo deste indicador é dado por:

$$GID = 1 - \frac{EC}{\hat{EC}} = 1 - \frac{\sum_{k=1}^K \sum_{i=1}^I \alpha_k^i \ln \frac{\alpha_k^i}{\alpha_k}}{\sum_{k=1}^K \sum_{i=1}^I \alpha_k \ln \frac{\alpha_k}{\alpha_k^i}} \quad (33)$$

Onde, \hat{EC} é a entropia cruzada estimada e EC é a entropia cruzada observada. A \hat{EC} mede a entropia entre a distribuição estimada e a distribuição agregada. Quando o valor da \hat{EC} é zero, a distância é nula e não existe heterogeneidade nos dados. A EC mede a entropia cruzada dos coeficientes da informação observada no nível desagregado e no nível agregado, dando a heterogeneidade de informação observada quando desagregada.

Segundo Howitt & Reynaud (2002), o GID deverá aumentar à medida que se ganha informação com a desagregação. O ganho é nulo quando a desagregação é uniforme e sem heterogeneidade. Quando o GID é 1, atinge o seu valor máximo, e indica que toda a informação foi recuperada.

No que diz respeito à validação e avaliação da utilidade prática do modelo, optou-se por realizar a comparação das estimativas com os dados observados, através do cálculo do desvio absoluto médio (DAM) por factor de produção ou item de custo (i) e por actividade agrícola (k):

$$DAM_k^i = \frac{\alpha_k^i - \alpha_k}{\alpha_k^i} \times 100 \quad (34)$$

Onde α_k^i representam os dados observados que são comparados com os resultados estimados.

4. INFORMAÇÃO DISPONÍVEL E IMPLEMENTAÇÃO EMPÍRICA DO MODELO

Neste capítulo apresentam-se as características da base de dados utilizada nesta dissertação. Esta base de dados, a rede de informação contabilística agrícola (RICA), rege-se pelos regulamentos comunitários onde se encontram espelhados os objectivos e normas a seguir para o seu bom desempenho na União Europeia (UE) e em Portugal. Começa por fazer-se uma descrição da RICA comunitária através dos campos de selecção da amostra, das definições da Tipologia, da estratificação e ponderação das empresas seleccionadas. De seguida aborda-se o problema da comparabilidade e controlo dessa base de dados. Apresentam-se ainda os dados da amostra RICA portuguesa com as suas especificidades e a orgânica de base relativa aos encargos da exploração e ao ficheiro de resultados. Por último, faz-se referência ao modo de selecção prévia das explorações agrícolas, das actividades e dos custos e descreve-se o processo de organização e de tratamento de dados da amostra RICA.

4.1 - Caracterização da RICA

A RICA é um instrumento criado para avaliar o rendimento das explorações agrícolas e os impactos da Política Agrícola Comum (PAC) (<http://ec.europa.eu/agriculture/rica>). É a única fonte harmonizada de dados microeconómicos do sector agrícola dentro da UE, ou seja nesta rede os princípios de contabilidade são os mesmos para todos os Estados-membros, regendo-se pelos mesmos conceitos e definições, através de regulamentos comunitários apropriados. A Comissão Europeia é a principal utilizadora de análises baseadas em dados da RICA, no entanto os dados agregados estão disponíveis e podem ser utilizados para estudos diversos.

A RICA teve início em 1965, quando o Regulamento 79/65 estabeleceu os conceitos e a base jurídica para a organização e desenvolvimento da rede: os Estados-membros da UE têm de fazer um levantamento anual, realizado pelos seus serviços operacionais, de informação qualificada para a UE. É função da RICA recolher todos os dados contabilísticos referentes a um ano, a partir de uma amostra de explorações agrícolas da UE, através de inquéritos nacionais.

A unidade responsável pela RICA (DGAgri) no âmbito da Comissão fornece uma visão geral anual da economia agrícola da UE com base em diversos indicadores económicos e financeiros: Análise da situação das explorações agrícolas da UE e sua evolução por Estado-Membro e por tipo de exploração nos últimos anos; análises sectoriais diversas (cereais, leite, porco, carne de bovino) a cada ano, com estudos de margens por unidade de produto e por rendimento agrícola; relatórios sobre a evolução do rendimento, a distribuição dos pagamentos directos e Zonas Desfavorecidas. Para além dos relatórios periódicos, a unidade responsável pela RICA executa também análises específicas, que são solicitados por diversas unidades e direcções da Comissão e outras instituições da União Europeia, especialmente unidades encarregadas da gestão dos mercados agrícolas, desenvolvimento rural, a avaliação das medidas da Política Agrícola Comum (PAC) e formulação de políticas. A procura por dados da RICA e pelas suas análises é contínua e crescente. As análises RICA desempenham realmente um papel essencial em todo o processo de reforma da PAC, ou seja, a avaliação das medidas da PAC e dos seus impactos. Os relatórios mais pertinentes podem ser encontradas no sítio <http://ec.europa.eu/agriculture/rica/>.

Seleção da amostra RICA

As explorações agrícolas são seleccionadas para participar da recolha de informação com base em planos de amostragem estabelecidos ao nível de cada região na UE. O levantamento não cobre todas as explorações agrícolas da UE, analisando-se apenas aquelas que, devido à sua dimensão, podem ser consideradas comercialmente activas³.

A RICA requer informação das explorações agrícolas acima de um determinado limiar mínimo de rendimento⁴, de modo a incluir a parte mais importante da agricultura da UE.

Essa parte implica, no mínimo, 90% da superfície agrícola utilizada e da margem bruta padrão total⁵. Actualmente, a amostra anual cobre aproximadamente de 80.000 explorações.

³ EA comercial é definida como uma exploração que seja grande o suficiente para proporcionar uma actividade principal ao agricultor e um nível de rendimento suficiente para sustentar sua família

⁴ Limiar determinado por cada estado membro e pela DGAgri.

⁵ Segundo a Tipologia Comunitária das Explorações Agrícolas

Para a UE-27, a RICA representa à volta de 6,4 milhões de explorações agrícolas (<http://ec.europa.eu/agriculture/rica>), de um total de cerca de 14 milhões de explorações agrícolas incluídas no Inquérito à Estrutura das Explorações Agrícolas 2007, última publicação do Eurostat.

A amostra deve reflectir a heterogeneidade da agricultura e para garantir essa diversidade, antes de se determinar a amostra de explorações agrícolas, é da competência da Agência de Ligação estratificar o campo de observação, definido de acordo com três segmentos/critérios: região RICA, dimensão económica e tipo de orientação. A estratificação é uma técnica estatística destinada a aumentar a eficiência da amostragem, minimizando o número necessário de explorações agrícolas para representar a diversidade no campo de observação.

Na RICA, o universo (o termo estatístico utilizado para definir o conjunto de unidades em estudo) de explorações agrícolas é constituído pelo conjunto das explorações que cumprem os seguintes requisitos:

- possuem uma área produtiva mínima de 1 hectare;
- apresentam menos de 1 hectare mas fazem a comercialização de uma parte da sua produção;
- apresentam menos de 1 hectare mas produzem mais que uma certa quantidade de produto (dentro de determinados níveis).

Quando as explorações agrícolas do universo não preenchem estes requisitos não são incluídas na amostra RICA. Portanto, no campo de observação, as explorações agrícolas da RICA são potencialmente explorações comerciais.

Na definição de campo de observação RICA, a Comissão segue as directrizes especificadas no Regulamento 79/65/CEE de 15 de Junho de 1965, e alterações posteriores, adoptando uma abordagem pragmática de incluir apenas as explorações agrícolas consideradas comerciais, sendo classificadas como tal as explorações agrícolas que ultrapassem um nível mínimo de dimensão económica.

A Tipologia Comunitária

A tipologia comunitária das explorações agrícolas é uma classificação uniforme das explorações da UE, com base na dimensão económica e no tipo de orientação técnico-económica.

Para se calcularem estes dois itens referentes à tipologia comunitária é necessário determinar as Margens Brutas Padrão (MBP).

A MBP de uma actividade vegetal ou animal é definida como o valor da produção bruta (PB) por hectare ou por animal menos os custos específicos proporcionais (CEP), ou seja dos custos variáveis referentes a essa produção: $MBP = PB - CEP$.

Depois de recolhidos anualmente, e numa base regional, os elementos necessários ao cálculo das MBP, estas são organizadas em médias de três anos para esbater as flutuações do ano e reflectir a diversidade da agricultura na União Europeia. O nível de detalhe deve ser de modo a garantir que os resultados dos inquéritos RICA e do Eurostat possam ser correctos e abrangentes. O conceito de Margem Bruta Padrão também é utilizado no Inquérito às Estruturas das Explorações Agrícolas, organizado pelo Eurostat para classificar as explorações inquiridas.

A Dimensão Económica

A dimensão económica (DE) das explorações agrícolas é a soma de todas as Margens Bruta Padrão das actividades de uma exploração agrícola. É expressa em termos de unidades de dimensão europeia (Ude)⁶ e apresenta-se dividida em classes.

São 6 as etapas para a determinação das classes de DE das explorações agrícolas: a) Identificar as actividades presentes na exploração; b) Determinar a situação da exploração agrícola (hectares ou número de animais); c) Multiplicar a dimensão das actividades pelas MBP correspondentes; d) Somar as diferentes margens brutas padrão das actividades; e) Obtém-se assim a margem bruta padrão total para a exploração agrícola; f) Determinar a dimensão económica da exploração agrícola, dividindo a margem bruta total pelo valor da Ude.

⁶ Ude=1200€

A Orientação Técnico-Económica

A Orientação Técnica Económica (OTE) das explorações agrícolas é determinada pelo seu tipo de sistema de produção. É definida pela contribuição relativa das diferentes actividades para a MBP total da exploração. Existem 9 níveis de orientação técnica, podendo ser mais desagregadas para melhor representar os sistemas de produção do Estado Membro.

A classificação das explorações agrícolas por tipo de agricultura baseia-se:

- a) Na definição das cinco principais classes de produção das explorações especializadas: 1-Culturas arvenses; 2- Horticultura (hortícolas e flores); 3- Culturas permanentes (pomares, olival e vinha); 4- Herbívoros (bovinos, equídeos, ovinos e caprinos); 5- Granívoros (suínos, aves e coelhos). No caso de Portugal, as plantas forrageiras tuberosas, plantas forrageiras, prados permanentes e pastagens pobres deixam de ser contabilizadas desde que haja herbívoros na exploração agrícola uma vez que se convencionou que o país apresenta défice forrageiro.
- b) Na importância das explorações mistas: 6- Policultura, de uma combinação de produtos vegetais; 7- Polipequária, de uma combinação de produtos de animais; 8- Mistas de uma combinação de culturas e de produtos de animais.
- c) Nas que não satisfazem qualquer das situações anteriores: 9- Explorações não classificadas.

Delimitação de campo de observação RICA

Como só as explorações agrícolas ditas comerciais podem entrar no campo de observação, é feita uma selecção daquelas que ultrapassam um determinado limiar de DE fixado pela Comissão para cada Estado-Membro. De seguida, as agências de Ligação desenvolvem os seus planos amostrais para a própria selecção da amostra e esses planos são apresentados ao Comité RICA para aprovação.

4.2 - Estratificação do campo de observação

Após a delimitação do campo de observação RICA, faz-se uso da estratificação devido à grande diversidade de agricultura, quer pela dimensão económica quer pelo tipo de produção: número de hectares da exploração agrícola que se concentram na produção agrícola, especialização em pecuária, ou explorações agrícolas com actividades mistas, apresentando diversos tipos de orientação.

Com apenas estes dois critérios, dimensão económica e tipo de orientação, o campo de observação de explorações agrícolas da União Europeia fica altamente heterogéneo. Para garantir que a amostra reflecte correctamente essa heterogeneidade, as Agências de Ligação estratificam o campo de observação antes da selecção da amostra. Se essa estratificação não for feita, há o risco de determinados tipos de explorações agrícolas não serem adequadamente representados. A Comissão faz uso desta técnica e utiliza três critérios para a estratificação: a região, dimensão económica e tipo de orientação. A obtenção da amostra da RICA pode ser representada como uma matriz tridimensional de células (Fig. 4.1), onde ao universo (*Universe*) é retirado o campo de observação (*Field of survey*), depois de ter sido estratificada por tipos de orientação técnica (*Types of farming*), por classes de dimensão económica (*Economic size classes*) e por regiões (*Regions*) conforme as condições impostas (*Clusters*). Por exemplo, se só se analisar as explorações com classes de DE superior a 2400 €, esta decisão vai determinar a amostra RICA (FADN Sample).

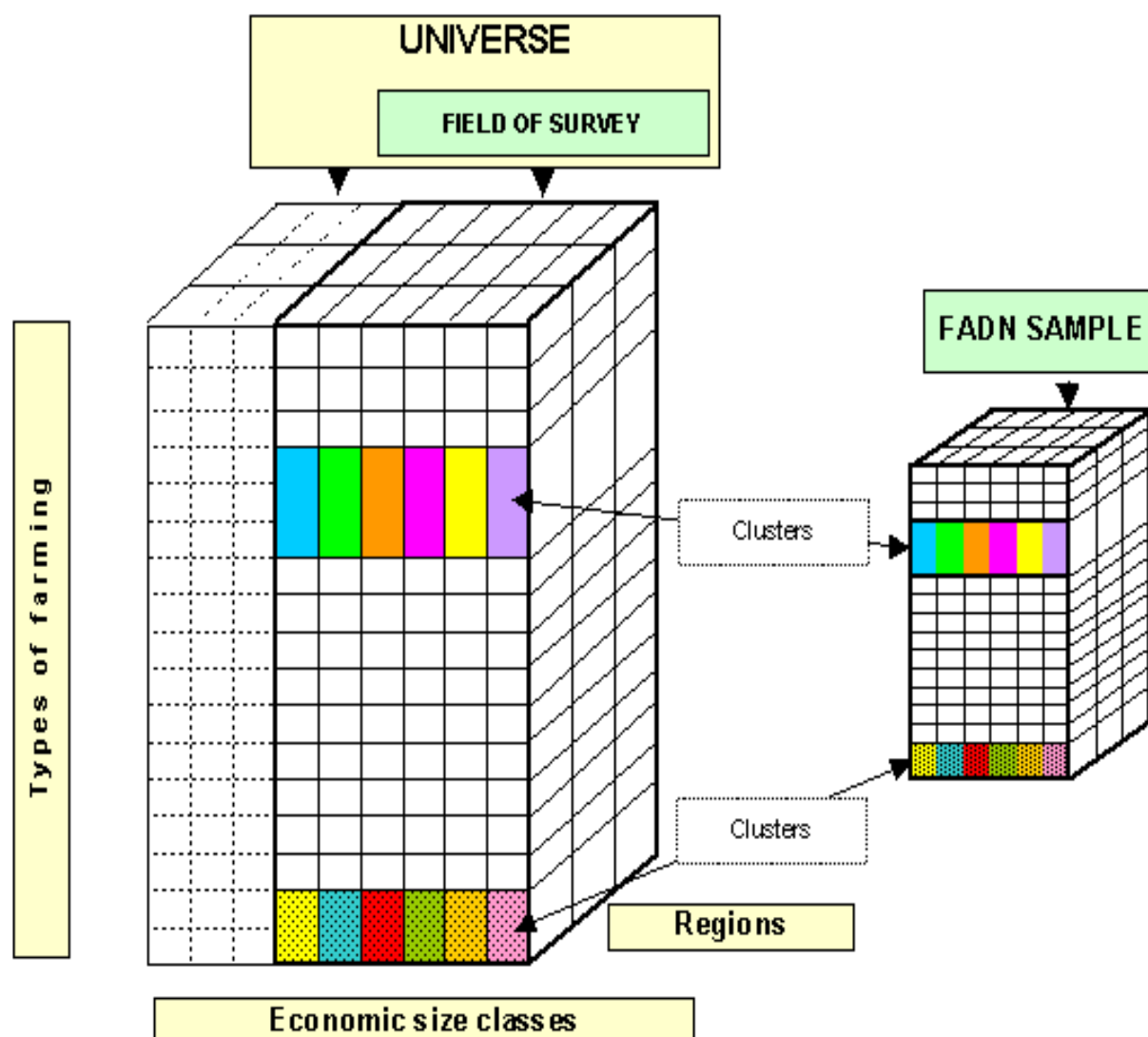
A base teórica utilizada para a estratificação é composta por um grupo numeroso de células de classificação (140 regiões da RICA x 72 tipos de orientação x 9 classes de dimensão económica = 90.720 células), para o ano do estudo (2008).

Por vezes, certas células não existem ou são muito raras em alguns Estados Membros (tipo de orientação x classes de dimensão económica).

A partir do conhecimento do campo de observação e do método de selecção, a Comissão, assistida pela Agência de Ligação dos Estados-membros, é capaz de julgar as células que podem estar vazias,

sem qualquer exploração a representar. As células que apresentam características semelhantes são agregadas e tratadas como se de uma única célula se tratasse.

Figura 4.1- Amostra RICA estratificada por classes de OTE, DE e Região



Fonte: <http://ec.europa.eu/agriculture/rica/>

4.2.1 - Ponderação

Um sistema de ponderação especial é usado no cálculo dos resultados da RICA, de modo a compensar as diferentes fracções de amostragem obtidas pela estratificação, nas diversas células calculadas. É baseado no princípio da "livre expansão": um peso calculado aplica-se a cada exploração da amostra (factor de extrapolação). O cálculo do peso individual das explorações na amostra e no campo de observação é feito de acordo com a estratificação segundo os critérios de região RICA, dimensão económica e tipo de orientação. O peso individual é igual à razão entre o número de explorações, da célula (região RICA x tipo de orientação x classe dimensão económica), no campo de observação e na amostra. Ex.: se houver 1000 explorações na célula estratificada, mas apenas 20 explorações colaboraram com a RICA então calcula-se o ponderador ($1000/20=50$) de modo a extrapolar o resultado dessas 20 explorações para representar as 1000 inicialmente pretendidas.

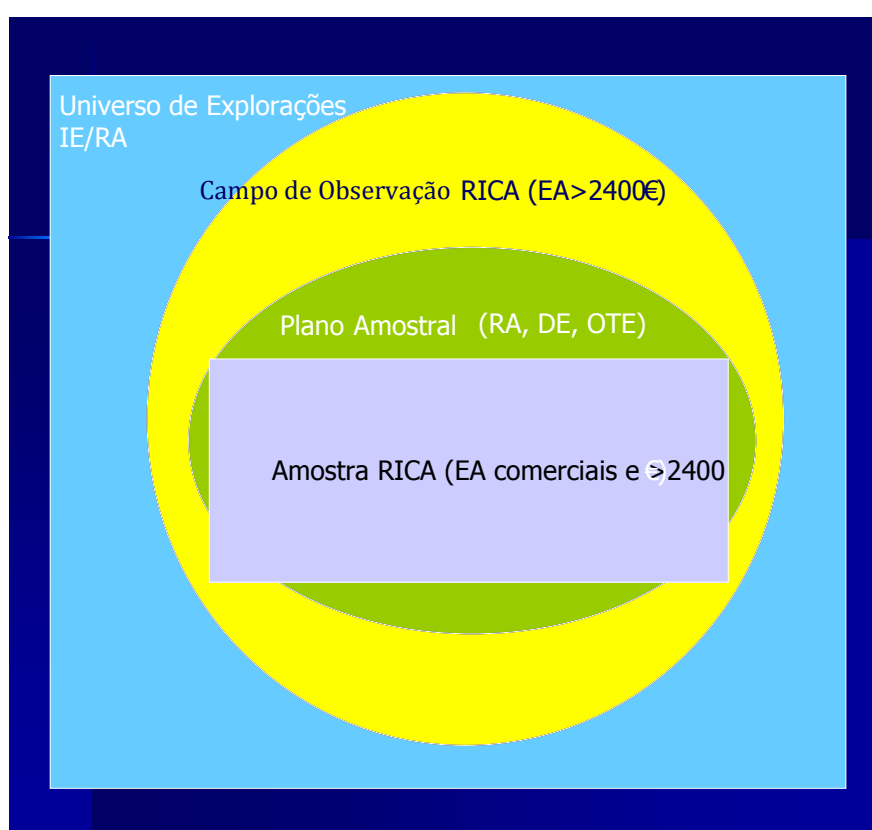
4.2.2 - Comparabilidade e Controlo

A comparabilidade da RICA a nível da União Europeia é suportada pela tipologia que foi criada para ser usada por diversos organismos e é uma metodologia que abrange os diversos tipos de agricultura. Esta tipologia, descrita na Decisão 85/377/CEE, de 7 de Junho de 1985, identifica os principais tipos de agricultura através da orientação técnica e da dimensão económica. Para o caso de Portugal, o Inquérito à Estrutura das Explorações Agrícolas 2007, apresenta um total de 275.090 explorações agrícolas, sendo o campo da RICA composto por 117.050 explorações agrícolas, com uma taxa de cobertura de 39%. Estes valores representam 77% da dimensão económica, 82% de SAU e 52% da UTA, valores acordados entre a Comissão e a Agência de Ligação.

A determinação do tamanho ideal da amostra nem sempre é conseguida. Fracções de amostragem variam de célula para célula. Em alguns Estados-membros, as agências de Ligação têm dados suficientes sobre a variabilidade das explorações agrícolas dentro do campo de observação para calcular fracções de amostragem ideal. Noutros casos, fracções de amostragem são definidos de

acordo com o número de explorações agrícolas na célula. Após a elaboração do plano de selecção ou amostral, as explorações agrícolas podem mudar de uma célula para outra, se houver uma mudança na sua dimensão económica ou no tipo de orientação técnica. Como a participação dos agricultores é voluntária, nem todos têm a contabilidade agrícola atempadamente disponível e com o decorrer dos anos a maior parte das explorações já foram inquiridas no ano anterior, pelo que a amostra não é aleatória e o plano amostral é meramente indicativo (Figura 4.2).

Figura 4.2- Representação do universo RICA



O universo das explorações é dado pelo Recenseamento Agrícola (RA) feito de 10 em 10 anos ou pelo Inquérito à Estrutura das Explorações Agrícolas (IE), cuja amostra é extrapolada para o universo, com as referências temporais de anos acabados em 3, 5 e 7. O campo de observação comporta as explorações agrícolas (EA) com valores DE superiores a 2400 € (caso português) e o

plano amostral está estratificado por região agrária (RA), dimensão económica (DE) e orientação técnico-económica (OTE) originando depois a amostra propriamente dita com apenas empresas comerciais.

Este procedimento é um exercício dinâmico, interactivo e concertado, de modo a tentar representar o mais possível a realidade de cada Estado-Membro. O exercício começa no ano $n-1$ com a apresentação do plano amostral com o número de explorações por Região, DE e OTE; no ano n é feita a recolha da informação que termina em meados do ano $n+1$, quando são feitas as validações e o apuramento de resultados para confrontar com o plano amostral. Se, por qualquer motivo alguma exploração não acabar o ano em exercício ou não atingir o limiar mínimo a cumprir, essa exploração é afastada da amostra RICA. A amostra do ano n fica finalizada no ano $n+1$, após a validação e fecho do ano por parte da DGAgri. Só então, são calculados os ponderadores a serem usados para representar o campo de observação.

A necessidade de controlo de qualidade leva a procedimentos exaustivos de controlo de qualidade seguidos pelas agências de Ligação e pela RICA, pois as decisões sobre a política agrícola na União Europeia devem ser baseadas em análises sólidas e precisas. Todos os controlos efectuados encontram-se bem detalhados nos regulamentos de suporte do RICA.

4.3 - A RICA em Portugal

Os países que sucessivamente foram aderindo à CEE e mais tarde UE ficaram desde logo submetidos ao Regulamento n.º 79/65/CEE que criou a RICA. Foi o caso de Portugal que, desde 1980, começou a organizar o sistema e fez o seu primeiro exercício contábil em 1983.

O plano de selecção e o estudo da representatividade das explorações agrícolas a participar na amostra têm de representar a realidade e a diversidade agrícolas das explorações agrícolas portuguesas.

O plano para a amostra RICA 2008 foi seleccionado com base na mais recente informação estrutural existente na altura, o IE 2005.

O campo de observação de RICA 2008 é constituído por todas as explorações com duas ou mais unidades de dimensão europeia e com base no Inquérito à Estrutura das Explorações Agrícolas de 2005 (IE 2005). Estas explorações representam cerca de 92,5% do total da MBP, 89,4% do total da Superfície Agrícola Utilizada (SAU) e 44% de explorações do universo de explorações agrícolas.

O estudo do campo de observação estratificado é feito também com recurso ao tratamento estatístico do IE 2005, o qual é cruzado com a informação estrutural das explorações agrícolas (SAU) e com os coeficientes económicos das MBP (DE e OTE).

O número total de explorações, para Portugal, para o ano de 2008, previa um acompanhamento de 2300 explorações, para as regiões agrárias Entre Douro e Minho e Trás-os-Montes (Norte), Beira Litoral e Interior (Centro), Ribatejo e Oeste (Lisboa e Vale do Tejo), Alentejo, Algarve, Açores e Madeira.

A Tabela 4.1 apresenta a distribuição dos sistemas de produção por classes de OTE em Portugal e o número de explorações agrícolas por classe de DE.

Tabela 4.1- Amostra RICA 2008: n.º de explorações agrícolas por classes de OTE e DE

	OTE \ DE (ude)	>2-8	>8-16	>16-40	>40-100	>100	Total EA
1	Culturas Arvenses	138	52	42	25	16	273
2	Horticultura Intensiva	72	44	40	26	21	203
3	Culturas Permanentes	297	126	85	52	20	580
4	Herbívoros	160	104	107	64	29	464
5	Granívoros	19	19	21	18	21	98
6	Policultura	185	65	41	23	18	332
7	Polípecuária	89	32	27	17	11	176
8	Mistas Cult.e Pecária	94	32	27	14	7	174
	Total por DE	1054	474	390	239	143	2300

Fonte: RICA, 2010

No Regulamento que rege a RICA existem 10 classes de Dimensão Económica mas devido à dimensão estrutural das explorações agrícolas no País e ao volume de rendimentos que as mesmas possam potencialmente auferir, agregaram-se as explorações em 5 classes.

Esta agregação permite facilitar a análise, ajustando-a à realidade, levando em conta a pequena participação de explorações agrícolas nas classes com maior rendimento (Tabela 4.2).

Tabela 4.2- Amostra RICA 2008: n.º de explorações agrícolas por classes de DE

Classe DE	Limites em UDE	Limites em €	Total EA
1	de 2 a menos de 8	2400 a <9600	1054
2	de 8 a menos de 16	9600 a <19200	474
3	de 16 a menos de 40	19200 a <48000	390
4	de 16 a menos de 40	48000 a <120000	239
5	igual ou maior do que 100	=>120000	143
Total			2300

Fonte: RICA, 2010

4.3.1 - Os encargos de exploração

Os encargos de exploração (em espécie e em natureza) correspondem ao consumo dos factores de produção. Para cada rubrica há um valor para o encargo observado.

A auto-utilização das produções efectuada na exploração, ligada directamente com a produção do exercício, nomeadamente no sector pecuário, também é considerada um encargo da exploração. Esta auto-utilização é avaliada a preço de mercado e não a preço de produção, quando o produto em questão se encontra disponível no mercado da região e quando se consegue determinar a quantidade de alimentos forrageiros consumidos.

Os encargos podem ser separados conforme o seu objectivo: produção vegetal, produção animal e auto-utilização.

Encargos com a produção vegetal

Na produção vegetal estão incluídos todos os encargos com a produção de culturas anuais, plurianuais, permanentes e de pastagens e forragens temporárias e permanentes, tais como:

- sementes e plantas compradas quer para a produção final quer para auto-utilização na produção animal, assim como os custos a ter com a selecção, limpeza e desinfectação. No caso das culturas

permanentes, a plantação é um investimento sendo tratado noutra ponto. Também são valorizadas as sementes e plantas auto-utilizadas ao preço do mercado;

- adubos e correctivos incluindo a terra tratada, a turfa e o estrume quando adquirido;

- produtos de protecção às culturas ou fitofármacos; são os produtos utilizados nas culturas contra as pragas, as doenças e as infestantes, assim como outras situações que podem causar danos nas culturas.

- outros encargos específicos das culturas, comportando as despesas directamente ligadas à produção: análises de solos e coberturas plásticas, conservação e transformação, embalagem e armazenamento, água de rega, encargos com a eliminação de resíduos de culturas e outros encargos não mencionados nas rubricas anteriores.

Encargos com a produção animal

Os encargos com a produção pecuária compreendem:

- alimentos comprados, nos quais se incluem os alimentos propriamente ditos, como as forragens (fardos de feno e rolos de silagem), e os concentrados (alimentos compostos, farinhas, cereais, etc.), os suplementos, os produtos lácteos, de preservação e conservação dos alimentos, as despesas tidas com o aluguer das pastagens colectivas, como as superfícies forrageiras (as camas e a palha incluem-se nos alimentos comprados);

- alimentos produzidos na exploração, produções auto-utilizadas, que abrangem todos os produtos produzidos e utilizados na exploração e que podem ser comercializáveis, como por exemplo o leite recolhido, excluindo o leite de amamentação. As camas e palhas só deverão ser consideradas se, na região e no ano em causa, houver comercialização desses produtos. São avaliados sempre ao preço do mercado, caso haja mercado para o produto em causa, e quando se consegue determinar a sua quantidade;

- gastos com os medicamentos e assistência veterinária;

- outros encargos específicos da pecuária, comportando as despesas directamente ligadas à produção: cobrição, inseminação artificial, limpeza e desinfectação, embalagem e armazenamento e outros encargos não mencionados nas outras rubricas.

Na RICA não se considera a substituição dos animais como encargo mas sim como variação de inventário, i.e. a diferença entre o valor de entrada e o valor de saída sendo colocado como produção do ano.

Todas as culturas são apresentadas individualmente ou agregadas conforme o regulamento em vigor, de modo a permitir comparabilidade da informação. O autoconsumo e os pagamentos em natureza são valorizados ao preço ao produtor.

4.3.2 - Instrumentos de trabalho

A informação é recolhida e analisada com base em alguns instrumentos de notação: ficha de exploração e resultados padrão.

Ficha de exploração

As características do agricultor são recolhidas num instrumento específico chamado Ficha de Exploração, onde a informação relativa a cada exploração agrícola da amostra fica armazenada.

As variáveis listadas referem-se às diversas características da exploração agrícola, tais como:

- a) Dados físicos e estruturais (localização, áreas para a agricultura, efectivo pecuário, mão de obra e outros);
- b) Dados económicos e financeiros: valor da produção das diferentes actividades, inventários, vendas e compras, custos de produção, activos, passivos, quotas de produção e subsídios (e os prémios relacionados com a aplicação de medidas de política da UE).

A informação contida na ficha de exploração é transmitida aos organismos de ligação pertencentes, na sua maioria ao Ministério onde se encontra o sector agrícola de cada Estado-Membro. A responsabilidade da recolha de dados é da Agência de Ligação de cada Estado-Membro e cada Agência de Ligação é guiada por um Comité RICA Nacional, que por sua vez comunica com as Estruturas da UE-DGAgri.

A ficha de exploração é uma base de dados com 1614 variáveis, para uma amostra de 2300 explorações agrícolas. Dado o grande volume de informação, para a manutenção de variáveis de Resultado e Indicadores, a informação é agregada e trabalhada num ficheiro de Resultados (FResult).

Resultados Padrão (FResult)

Os Resultados Padrão são um conjunto de estatísticas, calculados a partir de fichas de exploração e que são produzidos e publicados periodicamente pela Comissão. Eles descrevem, com um detalhe considerável, a situação económica dos agricultores, por diferentes grupos de explorações agrícolas. Como a RICA está directamente ligada com a agricultura, a sua amostra abrange toda a gama de actividades agrícolas existentes nas explorações agrícolas. Além disso, também se recolhem dados sobre as actividades não-agrícolas da agricultura, as quais, para além de utilizarem a mesma exploração, fornecem rendimento complementar à exploração agrícola.

Neste ficheiro a informação é tratada, calculando-se para cada exploração a respectiva OTE, DE, MPBTotal e Ponderadores, após um processo de regras de validação nacional e comunitário.

4.4 - Organização dos dados e construção do modelo

Para tentar responder ao problema da falta de informação relativa aos custos dos factores por actividade agrícola, esta dissertação tem como principal objectivo estimar os coeficientes de afectação desses custos às actividades agrícolas, a partir da base de dados da RICA.

Para o efeito desenvolveu-se um modelo de desagregação dos custos de factores de produção variáveis por actividade com base na teoria da máxima entropia. Este modelo é aplicado no contexto da região agrária do Alentejo, segundo os dados da RICA de 2008.

No entanto, para desenvolver este modelo é necessário proceder a uma selecção prévia das explorações agrícolas, das actividades e dos itens de custo considerados, bem como a organização e o tratamento dos dados da amostra RICA.

4.4.1 - Selecção das explorações, actividades e itens de custo

Uma análise preliminar dos dados ajuda a organizá-los de acordo com as categorias de custo específico e de rendimento, seleccionar explorações, actividades e custos a estudar.

Do ficheiro de resultados da RICA consta um extenso número de actividades, para cada uma das regiões de Portugal. Na impossibilidade de estudar todas as regiões e todas as actividades, optou-se por fazer uma selecção com base na distribuição das explorações por classes de OTE e Margem Bruta Total.

Numa primeira fase elegeu-se o Alentejo como região de estudo, devido ao facto de nesta região a dimensão física das explorações agrícolas ser relativamente elevada e de existir também alguma especialização das actividades produzidas.

A Tabela 4.3. apresenta a distribuição de explorações por classes de OTE e Margem Bruta Total (MBT), na amostra RICA de 2008 para a região Alentejo. Da sua análise, verifica-se que as OTEs que agregam mais explorações agrícolas são a OTE dos Herbívoros (47%), a OTE das Culturas Anuais (16%) e a OTE das Culturas Permanentes (15%).

Estas explorações representam 78% do total de explorações da amostra para o Alentejo e contribuem com 75% para a MBPT da região. Cada uma dessas OTE apresenta um número de explorações que permite tratar cada uma delas como uma amostra autónoma, o que permite obter resultados por OTE.

Tabela 4.3- N.º de explorações por classes de OTE e MBT no Alentejo

OTE	MBPT (€)	%	nº explorações	%
Culturas Arvenses	2312814	22	40	16
Horticultura Intensiva	305837	3	10	4
Culturas Permanentes	1077999	10	36	15
Herbívoros	4519264	43	115	47
Granívoros	279879	3	2	1
Policultura	372920	4	7	3
Polípecuária	533736	5	9	4
Mistas Culturas e Pecuária	1145318	11	28	11
Total	10547768	100	247	100

Fonte: RICA, 2010

Cada OTE inclui várias actividades ou produtos. A OTE 1 - Culturas Arvenses inclui o seguinte grupo culturas que totalizam 12 actividades: trigo, milho, arroz, outros cereais, proteaginosas, horto-industriais, oleaginosas, olivicultura, prados e pastagens permanentes, prados temporários, milho forragem e outras culturas forrageiras.

A OTE 3 - Culturas Permanentes é bastante menos diversificada e integra apenas 5 actividades: trigo, outros cereais, fruticultura, olivicultura e viticultura.

Por último a OTE 4 – Herbívoros, que teve de ser desagregada em 2 subgrupos, um composto apenas pela componente vegetal e outro pela componente animal. O subgrupo das actividades vegetais é composto por 8 actividades: trigo, outros cereais, olivicultura, prados e pastagens permanentes, pastagens pobres, prados temporários, milho forragem, outras culturas forrageiras. No subgrupo das actividades pecuárias consideram apenas 4 actividades que são: bovinos de carne, bovinos de leite, ovinos e caprinos.

Dadas as características dos dados, foi extraída uma amostra de conveniência de explorações agrícolas da base RICA 2008, por OTE. No sentido de tornar essa amostra mais equilibrada em termos de culturas, áreas e produções, desenvolveram-se alguns critérios, que passam também por apostar principalmente nas culturas com alguma área produtiva (culturas que sejam comercialmente activas) e nas explorações com alguma dimensão (para o Eurostat as empresas têm de ter mais de 5 ha de superfície agrícola utilizada). Também se analisaram e eliminaram as explorações que apresentavam problemas de ordem técnica como: a) as explorações que se

encontravam a fazer implantação de culturas permanentes (em período de carência de rendimentos), b) as empresas que apresentavam coeficientes económicos negativos, como as margens brutas, c) as explorações que tinham tido problemas de sanidade animal e consequente vazio sanitário e d) as empresas com uma especificidade determinante no sistema de produção que as tornavam diferentes, como a produção de cavalos (equídeos), com grandes e cuidadosos regimes alimentares e os suínos e as aves com uma vertente industrial.

A utilização dos critérios mencionados para a escolha das explorações agrícolas e actividades em cada uma das OTE deram aos seguintes procedimentos que foram levados a cabo na composição da amostra de explorações.

Para as explorações da OTE Culturas Arvenses e Culturas Permanentes retiraram-se os seguintes da base de dados para obter a composição da amostra:

- Actividades com áreas inferiores a 1% da área total das explorações envolvidas (cerca de 0,3% da área total);
- Empresas que apresentavam uma área total inferior a 1% da área total da OTE;

As explorações agrícolas com comportamentos únicos.

No caso da OTE Culturas Arvenses a aplicação destes procedimentos levou a que o número de explorações da amostra fosse de 29, i.e., cerca de 73% das 40 explorações agrícolas iniciais da RICA. No caso da OTE Culturas Permanentes, passou-se de 36 explorações agrícolas iniciais para 28, que representam mais de 90% da área total da OTE.

No caso dos Herbívoros, os procedimentos foram mais complexos, o que se compreende pelo facto de se ter dividido a amostra em dois subgrupos. Deste modo, retiram-se os casos que comportavam as seguintes situações:

- actividades com áreas inferiores a 1% da área total das empresas envolvidas e que estavam fora do contexto da OTE (cerca de 96 ha e 1% da área total).

- as actividades de equinos, suínos, aves e outros permanecendo apenas bovinos, ovinos e caprinos (cerca de 99 CN⁷ e 1% de CN Total).
- explorações agrícolas que apresentavam uma área total inferior a 1% da área total da OTE.
- explorações agrícolas com comportamentos únicos, como por exemplo explorações em que ocorreram uma grande mortalidade de animais e margens brutas negativas.

Das 116 explorações agrícolas iniciais, passou-se para 55 explorações agrícolas, que representam cerca mais de três quartos da área total da OTE e cerca de dois terços do total das CN. A constituição dos dois subgrupos levou a que cada um deles ficasse constituído por 54 explorações de herbívoros com componente vegetal e 55 explorações de herbívoros com componente animal.

De acordo com a metodologia usada para calcular as margens brutas padrão, os itens de custo considerados para as actividades vegetais foram sementes e plantas, fertilizantes, fitofármacos, outros encargos com as culturas. Para as actividades animais esses itens foram rações compradas, alimentos forrageiros comprados e produzidos na exploração e assistência veterinária e outros encargos com a pecuária.

O valor da margem bruta é definido em euros por actividade (vegetal ou animal). Para este trabalho a unidade usada para as culturas é o hectare e para os animais é cabeça normal (CN) e para isso foi necessário recalcular os valores de cabeça natural para a cabeça normal. A cabeça normal é uma unidade de medida padrão que permite a agregação das várias categorias do efectivo pecuário para fins de comparação. As cabeças normais são definidas com base nas necessidades de alimentação das diferentes categorias de animais e de acordo com os coeficientes descritos no Anexo I do Regulamento (CE) n.º 1200/2009 da Comissão.

⁷ Para calcular as CN para a agregação de bovinos com 2 ou mais anos, o valor aplicado foi de 1CN/cabeça.

4.4.2 - Organização dos dados

Em situações de omissão de produtividade física ou de preço de venda dos produtos, devido ao autoconsumo, optou-se pelo uso dos valores referenciados na região através de um regulamento que estima os valores da produção padrão (VPP). Este valor, tendo em conta o regulamento (CE) n.º 1242/2008, é estimado para todas as actividades, com a produtividade física e o preço de venda mais frequente observado na região e é da competência do GPP a sua organização.

Para validar o modelo de desagregação de custos por actividade é necessário dispor dos coeficientes dos custos observados e desagregados por factor de produção e produto. Para esse efeito utilizou-se a estrutura de custos de referência das actividades das MBP de 2004.

Nas culturas forrageiras e pastagens, a área utilizada é a que consta da ficha de exploração, uma vez que esta área não consta do Fresult. Também as produções destas actividades, sendo auto-utilizadas não apresentam valor na base de dados, pelo que o valor da sua produção teve de ser estimado através do Valor de Produção Padrão (VPP) 2008.

Em termos da implementação empírica do modelo de entropia desenvolvido no âmbito do Capítulo 3, foi necessário dispor de informação acerca dos custos por item, assim como de informação relativa ao produto bruto das explorações. Deste modo, organizou-se a informação disponível tendo em conta as variáveis explicadas, relativas aos custos por item de cada exploração da amostra e as variáveis explicativas que dizem respeito ao produto bruto de cada uma das explorações da amostra.

4.4.3 - Caracterização das empresas seleccionadas

De seguida apresenta-se uma breve caracterização das explorações agrícolas envolvidas no estudo por OTE seleccionado após os procedimentos já mencionados anteriormente.

Caracterização das Empresas de Culturas Arvenses- OTE 1

As características das 29 empresas seleccionadas na OTE 1 são apresentadas na Tabela 4.4.

Da sua análise verifica-se que as explorações têm uma área média de 97 ha de superfície agrícola, com um mínimo de 31 ha e um máximo de 309 ha de área total, mostrando uma grande variação na sua dimensão física.

Tabela 4.4 - Caracterização das explorações da OTE 1

	Média	Máximo	Mínimo	Var. Max. (%)	Var. Min. (%)	%
Valor da Produção (€/ha)						
Trigo	462	997	162	116	-65	8
Milho	2733	3459	1305	27	-52	8
Arroz	1585	2910	943	84	-41	6
Outros Cereais	389	1126	116	189	-70	11
Proteaginosas	146	305	45	109	-69	1
Horto-industriais	6286	14374	2808	129	-55	44
Oleaginosas	514	996	97	94	-81	6
Olivicultura	702	1565	56	123	-92	2
Prados pastagens permanentes	50	50	50	0	0	1
Prados temporários	390	390	390	0	0	8
Milho forragem	2375	2375	2375	0	0	5
Out. culturas forrageiras	190	190	190	0	0	1
Área (ha/EA)						
Trigo	15	64	0	333	-100	15
Milho	2	40	0	1685	-100	2
Arroz	3	34	0	1028	-100	3
Outros Cereais	22	103	0	363	-100	23
Proteaginosas	4	32	0	804	-100	4
Horto-industriais	6	45	0	693	-100	6
Oleaginosas	10	70	0	587	-100	11
Olivicultura	2	14	0	568	-100	2
Prados pastagens permanentes	12	159	0	1241	-100	12
Prados temporários	17	256	0	1394	-100	18
Milho forragem	2	50	0	2800	-100	2
Outras culturas forrageiras	2	34	0	1465	-100	2
Encargos (€/ha EA)						
Sementes	69	0	372	-100	439	8
Fertilizantes	102	0	825	-100	708	12
Fitofármacos	61	0	782	-100	1182	7
Outros encargos	51	0	247	-100	384	6
Margem Bruta (€/ha EA)	556	19	2206	-97	297	66

Fonte: RICA (2010)

Verifica-se que as actividades que apresentam valores de receita mais elevados são as culturas horto-industriais, milho grão e milho forragem. As produções médias alcançadas pelas empresas (€

de receita/ha produzidos), com os máximos e mínimos atingidos, mostram um intervalo relativamente grande em quase todas as actividades.

Os dados referentes às actividades forrageiras, por terem sido estimados, não mostram variação. As actividades que mais contribuem para receita global da OTE, são as horto-industriais (44%), os outros cereais (11%) e o trigo, milho e as pastagens temporárias (8%).

As actividades com maior área média são os outros cereais (22%), as pastagens temporárias (17%) e o trigo com 15% da área e também com o maior peso na superfície agrícola utilizada (SAU). São de salientar as grandes variações encontradas na área cultivada, principalmente no que se refere a forragens e pastagens.

Quando se analisam os encargos por hectare, também se verifica que os gastos são muito diferentes, sendo os intervalos com maior amplitude, os apresentados pelos fitofármacos e pelos fertilizantes. A margem bruta também apresenta alguma variabilidade mas bem menor que os encargos.

Caracterização das Empresas de Culturas Permanentes- OTE 3

As empresas seleccionadas na OTE 3 apresentam uma área média de 33 ha, com mínimo de 6 ha e um máximo de 149 ha.

As actividades que apresentam maior receita média por ha são a fruticultura seguida de viticultura, sendo a primeira a que apresenta menor variação máxima. O peso da contribuição de cada actividade no rendimento da OTE 3 é de 48% por parte da viticultura e 40% da olivicultura (Tabela 4.5).

No que se refere à área, a olivicultura é dominante (67%), seguida da vinha (20%) e o trigo com 8%. De notar que a fruticultura ocupa apenas 2% da área, no entanto apresenta os melhores resultados médios por ha (6076€), seguida da viticultura (3447 €). Quando se observam os encargos, verifica-se que as variações são grandes nos outros encargos e fitofármacos. A margem bruta apresenta a menor das variações mínimas.

Tabela 4.5 - Caracterização das explorações da OTE 3

	Média	Máximo	Mínimo	Var. Max. (%)	Var. Min. (%)	%
Valor de Produção(€)						
Trigo	547	744	353	36	-35	3
Outros Cereais	876	916	277	5	-68	2
Fruticultura	6076	7994	132	32	-98	7
Olivicultura	848	1715	58	102	-93	40
Viticultura	3447	6086	857	77	-75	48
Área (ha)						
Trigo	3	60	0	2103	-100	8
Outros Cereais	1	30	0	2619	-100	3
Fruticultura	1	11	0	1978	-100	2
Olivicultura	21	149	0	594	-100	67
Viticultura	6	30	0	368	-100	20
Encargos (€/ha EA)						
Sementes	17	30	0	77	-100	1
Fertilizantes	86	218	0	155	-100	6
Fitofármacos	103	715	0	593	-100	7
Outros encargos	41	314	0	672	-100	3
Margem Bruta (€/ha EA)	1180	4883	857	314	-27	83

Fonte: RICA (2010)

Caracterização das Empresas de Herbívoros por actividade agrícola- OTE 4

As empresas orientadas para herbívoros com actividade agrícola apresentam áreas diversas desde 60 ha a 413 ha, sendo a média de 163 ha.

A nível das actividades, as maiores receitas médias por ha são provenientes de milho forrageiro, trigo e pastagens temporárias (Tabela 4.6.).

No entanto, a maior contribuição em termos de receita para a OTE é dada pela actividade de prados temporários com 33% do valor, seguido de outras forragens e milho forrageiro com 21% e prados permanentes com 16%, num total de 91%, mostrando claramente a prioridade da empresa em fornecer alimentação aos seus animais.

A maior área média está concentrada nas pastagens permanentes que condiciona cerca de 52% da área total envolvida e as outras culturas forrageiras apresentam um peso de 21%.

Tabela 4.6 - Caracterização das explorações da OTE 4 componente vegetal

	Média	Máximo	Mínimo	Var. Max. (%)	Var. Min. (%)	%
Valor de Produção (€/ha)						
Trigo	396	646	224	63	-43	2
Outros Cereais	200	580	69	190	-65	6
Olivicultura	162	1503	16	828	-90	1
Prados pastagens permanentes	58	58	58	0	0	16
Pastagens pobres	19	19	19	0	0	0
Prados temporários	390	390	390	0	0	33
Milho forragem	2393	2393	2393	0	0	21
Outras culturas forrageiras	190	190	190	0	0	21
Área (ha/EA)						
Trigo	2	31	0	1450	-100	1
Outros Cereais	9	60	0	567	-100	6
Olivicultura	2	39	0	1850	-100	1
Prados pastagens permanentes	82	354	0	332	-100	52
Pastagens pobres	2	113	0	5550	-100	1
Prados temporários	24	265	0	1004	-100	16
Milho forragem	2	54	0	2600	-100	2
Outras culturas forrageiras	33	291	0	782	-100	21
Encargos (€/ha EA)						
Sementes	5	36	0	620	-100	1
Fertilizantes	13	60	0	362	-100	4
Fitofármacos	1	15	0	1400	-100	0
Outros encargos	1	30	0	2900	-100	0
Margem Bruta (€/ha EA)	321	1692	53	427	-83	94

Fonte: RICA (2010)

As empresas orientadas para herbívoros e com actividade vegetal também apresentam flutuações bastante elevadas nos encargos como sucede com os outros encargos e os fitofármacos, variações muito elevadas.

Caracterização das Empresas de Herbívoros por actividade animal

As empresas orientadas para herbívoros (componente animal) apresentam quantidades de cabeças normais diversas desde 6 CN a 269 CN, com uma média de 72 CN.

A nível das actividades, a maior receita média por CN é proveniente da actividade de bovinos de leite (Tabela 4.7.), sendo os bovinos de carne a actividade que mais contribui para as receitas da OTE, com cerca de 46%. São estes os que apresentam o maior efectivo médio (43CN), seguido

dos ovinos com 25CN. Por outro lado, essas mesmas espécies contribuem para com 64% e 34% do peso do efectivo da OTE.

Quando se observam os encargos por CN, qualquer que seja o encargo apresenta variações muito grandes.

Todos estes dados sobre as características das empresas das diversas amostras mostram que, apesar de se terem aplicado uma série de critérios para apurar e melhorar as amostras a envolver no estudo, continua a existir uma grande heterogeneidade nos valores das áreas, das receitas, dos encargos, que traduzem modos de produção e eficiências de utilização dos recursos bastante diferentes.

Tabela 4.7 - Caracterização das explorações da OTE 4 componente animal

	Média	Máximo	Mínimo	Var. Max. (%)	Var. Min. (%)	%
Valor de Produção (€/ha)						
Bovinos de carne	348	1001	70	188	-80	46
Bovinos de leite	3083	3083	3083	0	0	23
Ovinos	368	1398	110	280	-70	27
Caprinos	532	929	117	75	-78	3
Efectivo (CN/EA)						
Bovinos de carne	43	269	0	348	-100	60
Bovinos de leite	3	135	0	4400	-100	3
Ovinos	25	130	0	282	-100	34
Caprinos	2	56	0	1767	-100	3
Encargos (€/CN EA)						
Concentrados	120	907	0	656	-100	13
Forragens	16	142	0	788	-100	2
Outros encargos	57	270	0	374	-100	6
Margem Bruta (€/CN EA)	730	4332	67	493	-91	79

Fonte: RICA (2010)

5. RESULTADOS

Neste capítulo apresentam-se os resultados dos modelos de Tobit-MEG e Tobit-ECG aplicados a quatro amostras de explorações agrícolas da RICA no Alentejo para o ano de 2008. Como foi referido anteriormente, cada amostra corresponde a um grupo de explorações classificadas numa OTE específica. As OTE objecto de estudo incluem as culturas arvenses, as culturas permanentes e os herbívoros, que se encontram neste caso sub-divididos em duas OTE, a dos herbívoros – vegetal e a dos herbívoros – animal. As amostras consideradas são compostas por 29, 28, 54 e 55 explorações agrícolas, respectivamente.

Como já foi referido anteriormente, o modelo Tobit-ECG foi utilizado com o objectivo de melhorar as estimativas do modelo Tobit-MEG, fazendo uso de *prior* de informação que não estão contidos nas amostras objecto de estudo. Neste caso utilizou-se como *prior* de informação $(q_{i,k})$, a estrutura de custos utilizada pelo GPP do Ministério da Agricultura em Portugal calculadas para as margens brutas padrão do ano 2004 por actividade agrícola.

Deste modo, os resultados dos modelos de Tobit-MEG e Tobit-ECG são apresentados para cada uma das amostras de explorações agrícolas consideradas e são analisados, primeiro em termos dos indicadores de entropia normalizada e de capacidade de prescrição (pseudo- R^2), depois em termos dos ganhos de informação obtidos no processo de desagregação e por último em termos da sua validade prática através da comparação entre os resultados estimados e os dados observados para os coeficientes de afectação dos custos dos factores por actividade agrícola, através do cálculo do DAM.

Para obter os resultados dos modelos Tobit-MEG e Tobit-ECG foi utilizado o programa GAMS (General Algebraic Modeling System).

5.1. - Indicadores de Entropia Normalizada e Capacidade de Prescrição Pseudo-R²

Na Tabela 5.1. apresentam-se os resultados dos modelos MEG e ECG relativos aos indicadores da entropia normalizada e do pseudo-R² das explorações agrícolas da OTE das culturas arvenses.

O modelo de MEG apresenta para os parâmetros estimados um valor de entropia normalizada S_p de 0,708. Este valor é relativamente próximo da unidade, o que denota uma percentagem razoável de criação de nova informação no processo de desagregação de custos. No modelo de ECG, como seria de esperar, o valor da entropia normalizada S_p é menor, i.e., mais perto de zero. No que diz respeito à componente do erro, os valores obtidos para entropia normalizada S_w são em ambos os casos muito próximos da unidade (0,981 e 0,991).

Tabela 5.1.- Resultados de Entropia normalizada e pseudo-R² para os modelos de MEG e ECG das explorações agrícolas da OTE culturas arvenses

	MEG	ECG
Entropia Normalizada	0,708	0,307
Entropia Normalizada do erro	0,981	0,991
Pseudo R²		
Sementes e Plantas	0,867	0,928
Fertilizantes	0,822	0,864
Pesticidas	0,753	0,803
Outros Custos	0,773	0,800
Margem Bruta	0,916	0,968

Fonte: Resultados do Modelo

Os valores obtidos para o pseudo-R², como referido anteriormente, mostram a capacidade explicativa do modelo. No MEG, os custos com sementes e plantas e com fertilizantes apresentam valores de 0,867 e 0,822, enquanto os gastos com os pesticidas e outros custos registam valores de 0,753 e 0,773. Todos esses valores, assim como o valor obtido para a margem bruta (0,916) estão perto da unidade, o que evidencia a existência de um forte poder explicativo que permite relacionar a distribuição dos custos dos factores de produção com a distribuição do produto bruto nas actividades das explorações agrícolas da OTE das culturas arvenses. No modelo da ECG, esse poder explicativo é ainda maior, uma vez que os valores do pseudo-R² são ainda superiores.

Da observação da Tabela 5.2., o modelo MEG relativo às Culturas Permanentes apresenta uma entropia normalizada para o vector S_p de 0,450, com uma criação de informação mais contida e para o vector de suporte do erro, S_w , o valor de 0,974, mais perto da unidade.

Nos valores obtidos pelo pseudo- R^2 , os custos com os pesticidas e os fertilizantes apresentam os valores mais próximos da unidade, com 0,893 e 0,845 respectivamente, enquanto os encargos com as sementes e plantas são 0,636 e outros custos apresentam valores de 0,293. Aquele grupo está bem mais perto da unidade. O item que apresenta valores mais próximos da unidade é a Margem Bruta (0,977).

Tabela 5.2.- Resultados de Entropia normalizada e pseudo- R^2 para os modelos de MEG e ECG das explorações agrícolas da OTE culturas permanentes

	MEG	ECG
Entropia Normalizada	0,450	0,253
Entropia Normalizada do erro	0,974	0,981
Pseudo R^2		
Sementes e Plantas	0,636	0,643
Fertilizantes	0,845	0,934
Pesticidas	0,893	0,920
Outros Custos	0,293	0,337
Margem Bruta	0,995	0,997

Fonte: Resultados do Modelo

No modelo da ECG, a entropia normalizada de S_p é mais pequena (0,253), portanto mais perto do zero, o que mostra que a incerteza de informação é bem pequena. A entropia normalizada do erro ruído também está muito perto da unidade (0,981).

Os valores do pseudo- R^2 dos custos de fertilizantes e pesticidas no modelo ECG passam para 0,934 e 0,920, aproximando-se mais da unidade. No caso dos custos com as sementes e plantas (0,643), os valores do pseudo- R^2 crescem muito pouco, apenas 0,007, e nos outros custos passam para 0,337. A margem bruta está com 0,997, praticamente na unidade.

O grupo Herbívoro na sua componente de custos agrícolas para actividades vegetais, apresentam para o modelo MEG uma entropia normalizada para o vector S_p de 0,489 e para o erro, S_w o

de 0,975. Nos valores obtidos para o pseudo- R^2 , e no mesmo modelo de MEG, os custos com as sementes e plantas e fertilizantes estão com 0,501 e 0,590, enquanto os pesticidas e outros custos apresentam valores de 0,229 e 0,231. A Margem Bruta atinge o valor de 0,987 que, quando comparado com a diversidade das outras variáveis, é bem melhor (Tabela 5.3).

Tabela 5.3.- Resultados de Entropia normalizada e pseudo- R^2 para os modelos de MEG e ECG das explorações agrícolas da OTE herbívoros-componente vegetal

	MEG	ECG
Entropia Normalizada	0,489	0,198
Entropia Normalizada do erro	0,975	0,980
Pseudo R^2		
Sementes e Plantas	0,501	0,479
Fertilizantes	0,590	0,582
Pesticidas	0,229	0,357
Outros Custos	0,231	0,239
Margem Bruta	0,987	0,983

Fonte: Resultados do Modelo

No modelo de ECG, a entropia normalizada dos coeficientes de probabilidade estimados S_p é mais pequena (0,198), mais perto do zero e a entropia normalizada do ruído está muito perto da unidade (0,980). Nos valores do pseudo- R^2 , os custos com os fertilizantes e as sementes e plantas passam para 0,582 e 0,479, mostrando uma pequena diminuição. Os outros custos crescem muito pouco, apenas 0,008, os pesticidas atingem 0,337 (crescem 0,218) e a margem bruta fica em 0,983.

O grupo Herbívoro na sua componente de custos direccionados para animais, apresenta no modelo de MEG (Tabela 5.4.), uma entropia normalizada para as probabilidades estimadas S_p de 0,755 e para o erro, S_w , o valor de 0,992. Os valores obtidos para o pseudo- R^2 , no MEG, para os custos com concentrados e com as forragens são 0,940 e 0,801, enquanto os outros custos apresentam valores de 0,654. A Margem Bruta atinge valores de 0,895, relativamente abaixo do valor atingido pelo concentrado.

Quando se passa para o modelo de ECG, há uma diminuição dos valores na entropia normalizada S_p e o valor da entropia normalizada do erro S_w está ainda mais perto da unidade, mantendo-

se as posições relativas das diversas actividades. Já a margem bruta apresenta um valor superior, 0,920.

Tabela 5.4.- Resultados de Entropia normalizada e pseudo-R² para os modelos de MEG e ECG das explorações agrícolas da OTE herbívoros-componente animal

	MEG	ECG
Entropia Normalizada	0,755	0,403
Entropia Normalizada do erro	0,992	0,993
Pseudo R²		
Concentrados	0,940	0,931
Forragens	0,801	0,784
Outros custos	0,654	0,654
Margem Bruta	0,895	0,920

Fonte: Resultados do Modelo

5.2 - Ganhos de informação desagregada

Nos ganhos de informação desagregada total com a utilização dos modelos (Tabela 5.5.), para as culturas arvenses, o indicador é de 0,871 para o MEG, que pode ser considerado de bom, mas ainda é melhor no ECG, com 0,999.

Tabela 5.5.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE culturas arvenses

	MEG	ECG
Ganhos de informação desagregada total	0,871	0,999
Ganhos de informação desagregada parcial		
Trigo	0,87	1,00
Milho	0,99	0,99
Arroz	0,98	0,99
O. cereais	0,98	1,00
Proteaginosas	0,82	1,00
Horto industriais	0,99	0,99
Oleaginosas	0,87	1,00
Olivicultura	0,83	1,00
Pastagens Permanentes	0,84	1,00
Prados Temporários	0,90	1,00
Milho forrageiro	0,91	1,00
Outras forragens	0,92	1,00

Fonte: Resultados do Modelo

O GID parcial evidencia na generalidade das actividades envolvidas, os ganhos que ocorrem parcialmente e como influenciam o GID.

No modelo de MEG, são visíveis 2 grupos com valores menos elevados mas mesmo assim bons: as proteaginosas, a olivicultura e a pastagem permanente com valores entre 0,82 a 0,84, e depois o trigo e as oleaginosas com 0,87.

No modelo de ECG, os ganhos de informação são quase totais e os que apresentam no MEG um valor menor acabam elevando esse valor quando se muda de modelo. O milho, o arroz e as hortícolas ficam com o valor de 0,99.

No caso do modelo de ECG, os ganhos de informação são superiores aos obtidos para o modelo de MEG e são quase totais, uma vez que o valor do indicador é na maior parte dos casos igual à unidade.

Tabela 5.6.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE culturas permanentes

	MEG	ECG
Ganhos de informação desagregada total	0,977	0,994
Ganhos de informação desagregada parcial		
Trigo	0,98	0,98
Outros cereais	0,88	0,90
Fruticultura	0,98	0,98
Olivicultura	0,99	1,00
Viticultura	0,99	0,99

Fonte: Resultados do Modelo

Para as culturas permanentes, os ganhos de informação total (Tabela 5.6.), são bons, o indicador é de 0,977, mas ainda é melhor no modelo de ECG com 0,994. Quando se analisa os GID parciais, apenas os outros cereais tem um comportamento algo diferente.

Nos ganhos de informação desagregada total (Tabela 5.7.), para os herbívoros na componente vegetal, o indicador é de 0,924 mas é superado pelo valor dado pelo modelo de ECG (0,992). As pastagens pobres e a olivicultura com ganhos relativamente menores no modelo de MEG, de 0,82

e 0,90, chegam a 1,0 no modelo de ECG. As restantes actividades ou estão com 1,0 ou nas redondezas de 0,98 e 0,99.

Tabela 5.7.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente vegetal

	MEG	ECG
Ganhos de informação desagregada total	0,924	0,992
Ganhos de informação desagregada parcial		
Trigo	0,97	0,99
O. cereais	0,97	0,98
Olivicultura	0,90	1,00
Pastagens Permanentes	0,95	0,98
Pastagens Pobres	0,82	1,00
Prados Temporários	1,00	0,99
Milho forragem	0,99	0,99
Outras forragens	0,99	1,00

Fonte: Resultados do Modelo

Ao se analisar os herbívoros na sua componente animal e os ganhos de informação desagregada total (Tabela 5.8.), com o modelo de ECG, os valores são muito superiores quando comparados com os do modelo MEG. Nos ganhos de informação parcial os valores de ovinos e de caprinos são, no modelo de MEG, 0,45 e 0,78, respectivamente, e atingem 1,00 no modelo de ECG.

Tabela 5.8.- Ganhos de informação dos modelos de MEG e ECG no processo de desagregação da OTE herbívoros componente animal

	MEG	ECG
Ganhos de informação desagregada total	0,814	0,995
Ganhos de informação desagregada parcial		
Bov.de carne	0,96	0,99
Bov.de leite	0,96	0,99
Ovinos	0,75	1,00
Caprinos	0,78	1,00

Fonte: Resultados do Modelo

5.3 - Validação das Estimativas dos Coeficientes de Afecção dos Custos

Nas Tabelas seguintes apresentam-se os resultados das estimativas dos coeficientes de afectação dos custos dos factores de produção às actividades agrícolas obtidos com os modelos de MEG e de

ECG, e procede-se à sua comparação com os dados observados através do cálculo do Desvio Absoluto Médio (DAM).

No caso do modelo de MEG das culturas arvenses, cujos resultados são apresentados no Tabela 5.9., verifica-se uma coerência generalizada dos valores estimados. No caso dos itens de custo relativos às sementes e plantas, o arroz, as proteaginosas e outras forragens apresentam valores de desvios elevados. Para os fertilizantes, têm-se o arroz, as horto-industriais, a olivicultura e ainda as oleaginosas com um valor bem elevado. Os pesticidas têm desvios elevados nos outros cereais e olivicultura. Os outros custos têm desvios consideráveis nas diversas actividades, talvez devido a ser um grupo de custos muito diversificado e ter também adesões diferentes. As margens brutas apresentam desvios entre 17% nas horto-industriais e 77% nas proteaginosas. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (44%), nos fertilizantes (45%), nos pesticidas (67%), nos outros custos (25%) e margens brutas (33%).

Tabela 5.9.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas arvenses

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,177	60,9	0,206	17,6	0,154	18,9	0,185	0,0	0,279	36,6
Milho	0,162	32,5	0,236	24,2	0,139	39,0	0,184	67,3	0,278	20,6
Arroz	0,190	216,7	0,210	133,3	0,237	5,2	0,131	0,8	0,232	51,7
O. cereais	0,153	17,7	0,225	2,2	0,143	615,0	0,180	100,0	0,299	44,6
Proteaginosas	0,198	120,0	0,200	0,0	0,195	0,0	0,201	0,0	0,206	77,4
Horto indust.	0,078	13,3	0,152	204,0	0,086	4,4	0,064	220,0	0,621	17,2
Oleaginosas	0,164	2,5	0,231	670,0	0,175	0,0	0,154	413,3	0,275	64,7
Olivicultura	0,196	0,0	0,204	308,0	0,196	553,3	0,193	0,0	0,212	76,7
P.Permanentes	0,200	0,0	0,200	81,8	0,190	0,0	0,198	120,0	0,213	73,4
P.Temporários	0,127	0,0	0,199	5,2	0,119	0,0	0,154	3,8	0,401	36,3
M. forrageiro	0,124	55,0	0,211	75,8	0,130	0,0	0,135	237,5	0,399	47,5
Out. forragens	0,196	292,0	0,202	34,7	0,195	0,0	0,197	0,0	0,210	72,7

Fonte: Resultados do Modelo

A análise descritiva dos DAM no modelo de MEG para as culturas arvenses, mostra que existem medianas que se comportam de uma forma aceitável, uma variação muito elevada entre os máximos e os mínimos originando uma média mais elevada, consequentemente um desvio padrão e coeficiente de variação também elevados (Tabela A 5.1.).

Na observação dos desvios no ECG para as culturas arvenses (Tabela 5.10.), percebe-se um melhor relacionamento dos coeficientes de afectação e os desvios com os valores reais, estão relativamente próximos. As situações mais elevadas são os fertilizantes e outros custos nas horto-industriais, devido a esta actividade ser um agregado de culturas, os pesticidas no milho (que agrega o milho grão, quer de sequeiro quer de regadio, híbrido e regional), outros cereais e outras forragens, que também são grandes agregados de culturas com comportamentos diferentes em produção.

Tabela 5.10.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas arvenses

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,113	28,0	0,23	4,4	0,163	37,0	0,000	0,0	0,494	3,8
Milho	0,166	42,5	0,255	72,1	0,149	203,0	0,180	63,6	0,249	17,9
Arroz	0,082	11,9	0,116	58,3	0,267	93,1	0,074	43,1	0,461	2,6
O. cereais	0,100	36,4	0,238	32,1	0,029	180,1	0,140	55,6	0,494	2,8
Proteaginosas	0,091	28,2	0,000	0,0	0,000	0,0	0,000	0,0	0,909	4,1
Horto indust.	0,082	31,8	0,106	137,4	0,083	58,1	0,053	165,0	0,676	7,9
Oleaginosas	0,125	39,9	0,038	66,7	0,000	0,0	0,033	10,0	0,803	13,2
Olivicultura	0,000	0,0	0,051	6,3	0,031	61,6	0,000	0,0	0,918	0,6
P. Permanentes	0,000	0,0	0,104	19,8	0,000	0,0	0,084	6,7	0,811	11,8
P. Temporários	0,000	0,0	0,121	24,6	0,000	0,0	0,099	38,1	0,780	45,5
M. forrageiro	0,046	56,8	0,089	9,9	0,000	0,0	0,029	27,5	0,836	17,2
Out. forragens	0,049	22,0	0,149	15,3	0,01	119,8	0,020	0,0	0,772	2,8

Fonte: Resultados do Modelo

Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (78%), nos fertilizantes (64%), nos pesticidas (14%), nos outros custos (50%) e margens brutas (92%).

A análise descritiva do DAM das culturas arvenses em ECG apresenta medianas aceitáveis, com desvios reduzidos; a variação de máximos e mínimas são bem menores, os desvios da média estão relativamente perto do valor da mediana e consequentemente os desvios padrão e coeficientes de variação são bem menores (Tabela A 5.2.).

No caso das explorações da OTE Culturas Permanentes, quando se observa os coeficientes de afectação e o DAM (Tabela 5.11.), no modelo de MEG, apenas os pesticidas nos outros cereais apresentam um valor muito elevado. Por outro lado, as margens brutas também apresentam desvios elevados para o trigo e outros cereais. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nos fertilizantes (60%), nos pesticidas (50%) e margens brutas (20%).

Tabela 5.11.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE culturas permanentes

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,143	58,7	0,301	1,1	0,241	19,6	0,017	0,0	0,297	154,7
Out. cereais	0,062	74,6	0,232	45,0	0,181	1370,1	0,047	90,6	0,478	478,4
Fruticultura	0,001	0,0	0,118	15,8	0,050	0,0	0,001	99,6	0,831	97,1
Olivicultura	0,003	0,0	0,114	32,5	0,070	0,0	0,007	93,7	0,806	28,0
Viticultura	0,010	0,0	0,037	77,7	0,083	0,0	0,016	59,3	0,854	58,6

Fonte: Resultados do Modelo

No caso das culturas permanentes, a análise descritiva do DAM no modelo MEG, revela que a mediana e a média têm valores algo diferentes, sendo nos outros custos a mediana superior à média. Os desvios máximos são de um modo geral inferiores aos que se obtiveram na OTE Culturas Arvenses, com excepção dos pesticidas que também apresentam desvios e coeficientes de variação elevados (Tabela A 5.3.).

Ao se mudar para o modelo de ECG das culturas permanentes (Tabela 5.12.), os coeficientes de afectação e o DAM dos fertilizantes na olivicultura têm um desvio considerável e de novo os pesticidas nos outros cereais apresentam um valor muito elevado seguido do trigo. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (50%), nos fertilizantes (20%), nos outros custos (100%) e margens brutas (80%).

Tabela 5.12.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE culturas permanentes

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,128	18,5	0,065	73,0	0,311	161,4	0,000	0,0	0,496	4,2
Out. cereais	0,040	74,6	0,131	27,3	0,126	1117,2	0,026	1,0	0,677	40,9
Fruticultura	0,000	0,0	0,101	85,2	0,000	0,0	0,000	2,0	0,898	6,0
Olivicultura	0,001	0,0	0,121	152,3	0,054	0,0	0,007	3,0	0,817	11,5
Viticultura	0,008	0,0	0,028	46,8	0,064	0,0	0,016	4,0	0,884	0,1

Fonte: Resultados do Modelo

A análise descritiva para as culturas permanentes no modelo ECG mostra que a mediana é muito aceitável, excepção para os fertilizantes e os pesticidas que apresentam máximos e médias muito grandes e têm desvios padrão e coeficientes de variação elevados (Tabela A 5.4.).

Nos coeficientes de afectação e desvios do DAM nos herbívoros na sua componente vegetal (Tabela 5.13.), no modelo de MEG, os fertilizantes apresentam desvios grandes e distribuídos pelas diversas actividades, assim como os outros custos nas actividades de outros cereais, pastagens permanentes e milho forragem e os pesticidas, nos outros cereais e olivicultura.

Tabela 5.13.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente vegetal

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,103	58,5	0,333	311,1	0,095	32,6	0,029	0,0	0,439	38,5
O. cereais	0,103	37,3	0,307	338,6	0,062	342,9	0,022	266,7	0,505	39,5
Olivicultura	0,142	0,0	0,293	2563,6	0,151	843,8	0,076	0,0	0,339	65,2
P. Permanente	0,070	0,0	0,222	753,8	0,046	0,0	0,012	140,0	0,651	32,8
Pastagens Pobres	0,203	0,0	0,209	0,0	0,200	0,0	0,176	0,0	0,211	78,9
P. Temporária	0,002	0,0	0,047	24,2	0,003	0,0	0,001	94,1	0,947	2,2
Mil. Forragem	0,022	42,1	0,070	141,4	0,018	0,0	0,009	350,0	0,880	5,5
Out. forragens	0,055	161,9	0,144	300,0	0,014	250,0	0,017	1600,0	0,771	17,7

Fonte: Resultados do Modelo

Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (25%), nos fertilizantes (14%), nos pesticidas (25%) e margens brutas (75%).

Na análise descritiva do DAM no modelo de MEG para as culturas da OTE herbívoros, os valores da mediana flutuam dentro de limites aceitáveis a pouco aceitáveis, donde para este grupo de actividades os coeficientes de variação são muito elevados (Tabela A 5.5.).

No caso do modelo de ECG (Tabela 5.14.), na OTE herbívoros na sua componente vegetal, os valores estão dentro de limites aceitáveis. Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nas sementes e plantas (25%), nos fertilizantes (57%), nos pesticidas (50%), nos outros custos (14%) e margens brutas (88%).

A análise descritiva do DAM para o modelo de ECG nos herbívoros na sua componente vegetal, mostra valores aceitáveis da mediana. Como os máximos não tem desvios muito elevados, a média está relativamente perto da mediana, os desvios padrão são relativamente baixos. No entanto, os coeficientes de variação são relativamente mais elevados (Tabela A 5.6.).

Tabela 5.14.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente vegetal

	Semente/Plantas		Fertilizantes		Pesticidas		Outros Custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM	α	DAM
Trigo	0,068	56,7	0,265	10,2	0,058	51,3	0,004	47,3	0,605	27,1
O. cereais	0,082	47,9	0,226	25,5	0,009	13,1	0,014	91,9	0,669	39,3
Olivicultura	0,000	0,0	0,050	4,3	0,015	21,8	0,006	37,5	0,930	0,7
P. Permanente	0,028	0,0	0,146	68,1	0,019	0,0	0,010	94,7	0,796	9,7
Pastagens Pobres	0,000	0,0	0,000	0,0	0,000	0,0	0,000	0,0	1,000	0,0
P. Temporária	0,000	0,0	0,022	86,3	0,000	0,0	0,000	99,9	0,978	82,4
Mil. Forragem	0,022	79,4	0,051	48,4	0,017	0,0	0,009	88,9	0,901	26,3
O. Forragens	0,049	22,0	0,088	31,9	0,001	87,3	0,017	67,8	0,845	12,6

Fonte: Resultados do Modelo

Ao se analisar os coeficientes e os desvios no modelo de MEG, para a componente animal da OTE Herbívoros, os bovinos apresentam desvios acentuados em todos os tipos de custos e os caprinos apenas nas forragens (Tabela 5.15.). Os desvios dos diversos custos com valores aceitáveis abaixo de 40% encontram-se apenas no item relativo a outros custos (25%).

Tabela 5.15.- Coeficientes de afectação (α) e DAM obtidos com o modelo MEG para as explorações agrícolas da OTE herbívoros-componente animal

	Concentrados		Forragens		Outros custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM
Bov. de carne	0,332	149,6	0,062	264,7	0,204	264,3	0,402	49,4
Bov. de leite	0,411	152,1	0,056	460,0	0,138	160,4	0,395	49,0
Ovinos	0,328	0,0	0,054	68,8	0,214	22,3	0,403	49,2
Caprinos	0,274	0,0	0,176	309,3	0,281	72,4	0,270	66,0

Fonte: Resultados do Modelo

A análise descritiva em MEG para a OTE Herbívoros na sua componente animal, salienta medianas e médias muito próximas e os desvios padrão algo elevados assim como os coeficientes de variação relativamente elevados, com excepção da margem bruta (Tabela A 5.7.).

Quando se passa para o modelo de ECG, os valores encontrados estão mais normalizados, só os outros custos continuam elevados para os bovinos (Tabela 5.16.). Os desvios dos custos com valores aceitáveis abaixo de 40% encontram-se nos custos relativos a concentrados (50%), às forragens (25%), aos outros custos (50%) e às margens brutas (100%).

Tabela 5.16.- Coeficientes de afectação (α) e DAM obtidos com o modelo ECG para as explorações agrícolas da OTE herbívoros-componente animal

	Concentrados		Forragens		Outros custos		Margem Bruta	
	α	DAM	α	DAM	α	DAM	α	DAM
Bov. de carne	0,254	16,0	0,046	55,9	0,124	220,9	0,577	9,6
Bov. de leite	0,403	48,9	0,052	17,7	0,116	211,5	0,429	31,8
Ovinos	0,000	0,0	0,034	83,0	0,163	27,1	0,802	19,4
Caprinos	0,000	0,0	0,146	42,5	0,156	39,3	0,698	10,1

Fonte: Resultados do Modelo

A análise descritiva do DAM no modelo de ECG, mostra que também neste caso a média e a mediana estão muito próximas e o máximo atinge um valor elevado nos outros custos, e consequentemente se traduz num aumento do desvio padrão. O coeficiente de variação apresenta também uma grande oscilação de valores (Tabela A 5.8.).

5.4 - Discussão de resultados

De um modo geral, os valores obtidos com a intervenção de indicadores de qualidade, precisão e estimação têm comportamentos muito diferentes, apesar de, em alguns casos serem muito semelhantes aos que são apresentados em diversos trabalhos como Fragoso *et al.*, (2011), Peeters & Surry (2002), nomeadamente, para as entropias normalizadas e para o pseudo- R^2 .

Também os valores apresentados pelos GID são relativamente semelhantes aos apresentados por Fragoso & Carvalho (2011), Xavier *et al.*, (2010) e Howitt & Reynaud (2003).

Os valores referentes aos coeficientes de afectação e aos DAM também são relativamente semelhantes aos apresentados por Fragoso & Carvalho (2011).

As principais dificuldades sentidas foram tal como referem Peeters & Surry (2002), problemas com as categorias heterogéneas como as outras oleaginosas e outros cereais onde se encontram grupos de culturas. Por outro lado, as variações ou melhor a heterogeneidade entre as empresas e, nomeadamente, a variabilidade entre os custos de factores também levaram a que existisse alguma disparidade entre os valores estimados e os valores reais observados, como aconteceu com Fragoso & Carvalho (2011).

Não obstante, os mesmos autores avançam com alguns valores para analisar o DAM na calibração de modelos, considerando que o modelo se encontra bem validado para valores inferiores a 15%. No entanto, um DAM superior a 15% pode ser considerado razoável, mas mostra que o modelo precisa de ser melhorado. Howitt & Msangi (2006) consideram aceitáveis valores de DAM entre 1.5 % e 40.9%.

6. CONCLUSÕES E RECOMENDAÇÕES

A falta de informação desagregada relativamente a factores de produção é um problema generalizado na área das ciências agrárias. Para se ter esse detalhe recorre-se à utilização de inquéritos directos e pessoais junto do agricultor, que são dispendiosos e demorados.

A recolha de dados ao nível da exploração é difícil e onerosa, com recurso a amostras pequenas, pelo que o uso de técnicas e métodos alternativos pode ser uma experiência válida para estimar os custos variáveis de produção unitários da exploração a um custo significativamente mais baixo.

A base de dados da Rede de Informação de Contabilidades Agrícolas (RICA) é uma fonte de informação muito útil para estudar diversos aspectos da política agrícola. A informação sobre as explorações constantes das bases de dados da RICA respeita ao total dos custos por tipo factores, não sendo desagregada por actividade. Para além disso, apesar de haver uma tendência para a especialização, a maioria das explorações agrícolas são ainda multiproduto.

A obtenção dos custos variáveis unitários de produção por unidade de actividade é extremamente importante quer para a gestão da exploração quer na perspectiva da análise da política agrícola. Daí o propósito deste estudo em desenvolver um modelo que permitisse obter, a partir da base de dados RICA, os encargos de produção desagregados por actividade. Para o efeito, utilizou-se uma metodologia baseada na teoria da Máxima Entropia, para estimar os coeficientes de afectação tendo em vista a desagregação do custo dos factores de produção ao nível da exploração agrícola.

Outros métodos poderiam ter sido utilizados, mas devido à dimensão das várias amostras e pela tendência de vários autores, decidiu-se testar e desenvolver uma variante Tobit de um modelo de Máxima Entropia Generalizada e utilizar a sua formulação de Mínima Entropia Cruzada para melhorar os resultados. Entre as principais vantagens da opção por este método está a forma como é possível ultrapassar dificuldades relacionadas com a incorporação de informação, flexibilidade e facilidade no manuseamento e também a existência de valores nulos na amostra.

Neste trabalho utilizou-se a informação proveniente da base da RICA 2008 para a região Alentejo, da qual se seleccionaram 4 amostras com base no peso relativo de cada Orientação Técnico-Económica na margem bruta padrão da região. Das amostras retiradas da base e após um processo de consistência de informação, reuniram-se 4 grupos de empresas agrícolas com 29, 29, 54 e 55 explorações, que correspondem a conjuntos de explorações agrícolas com Orientação Técnico-Económica em culturas arvenses, culturas permanentes, herbívoros- componente vegetal e herbívoros -componente animal, respectivamente.

Tendo em conta as limitações dos dados recolhidos, a utilização desta metodologia mostrou ser uma opção adequada ao problema da estimação dos coeficientes de afectação dos custos dos factores de produção num contexto de informação incompleta, uma vez que os resultados obtidos estão, na generalidade dos casos, dentro dos limites impostos pelos diversos critérios de qualidade e pelos valores reais observados. É de salientar que os resultados obtidos de um modo geral estão alinhados com os de outros autores que trataram também o problema da estimação de coeficientes de afectação de custos. Os estimadores de entropia apresentam valores que, na maioria dos casos, se podem considerar bons tendo em conta as suas propriedades estatísticas e econométricas. No entanto, os resultados dos modelos de Mínima Entropia Cruzada apresentam uma maior aderência à realidade do que os obtidos com os modelos de Máxima Entropia Generalizada.

Apesar do método da entropia apresentar resultados muito estimulantes no que diz respeito à obtenção de coeficientes para a estimação de custos em contexto de informação incompleta, registaram-se ainda alguns desvios nos resultados, que devem ser melhorados no futuro.

Ao analisarem-se detalhadamente as características das explorações verifica-se que há alguns pontos que eventualmente poderão ser melhorados. Na Orientação Técnico-Económica Culturas Arvenses, deveriam estar apenas culturas arvenses e não culturas horto-industriais, culturas permanentes, as forragens anuais e permanentes, que são muito diferentes em área, em tecnologia/sistema de produção e rendimento e por conseguinte também em termos dos custos de produção. As áreas das explorações estão inseridas num intervalo muito grande, entre 31 e 309 ha. O mesmo sucede com as receitas deste grupo, sendo as mais elevadas nas explorações de horto-

industriais e as mais baixas nas culturas e prados permanentes. O produto bruto mostra as diferenças acentuadas que existem na mesma actividade, sobressaindo o grau de intensificação da agricultura e consequentemente o uso maior ou menor de factores de produção como sementes e plantas, fertilizantes, pesticidas e outros encargos, com reflexo directo na margem bruta. Nas outras duas classes de Orientação Técnico-Económica consideradas para o estudo, permanecem também as mesmas dificuldades que foram referidas para as Culturas Arvenses.

Com estas recomendações, considera-se que num futuro próximo, a aplicação das novas formas de abordagem podem fomentar um novo desenvolvimento dos modelos para a verificação de novos dados e a previsão de melhores resultados. Por outro lado, a metodologia da entropia apresentou-se como um instrumento de trabalho flexível e robusto para ajudar a compensar falhas de informação, como é o caso dos custos dos factores de produção desagregados por actividade agrícola, não sendo também a sua implementação difícil.

BIBLIOGRAFIA

Arfini, F.; Capelli, M.G.; Cesaro, L.; Donati, M.; Marongiu, S. – “Methodology to assess the farm production costs using PMP farm models”. In *FACEPA*. 2009.

Arfini, F.; Donati, M.; Paris, Q. – “A national PMP model for policy evaluation in agriculture using micro data and administrative information”. Contributed paper to the international conference: *Agricultural policy reform and the WTO: Where are we heading?*, Capri, June 2003.

Barker, T.; Pesaran, M. – “Disaggregation in Econometric Modeling– An Introduction”. In *Disaggregation in Econometric Modeling*. London. T. Barker and M. Pesaran editors. 1990.

Bewley, R. *Allocation Model: Specification, Estimation and Applications*. Cambridge: Ballinger Publishing Company. 1986.

Campbel, R.; Hill, R. – “Imposing parameter inequality restrictions using the principle of maximum entropy”. *Journal of Statistical Computation and Simulation*, 76, pp. 985-1000. 2006.

Campbel, R.; Hill, R. – “A Monte Carlo study of the effect of design characteristics of the inequality restricted maximum entropy estimator”. *Review of applied Economics*, 1, pp. 53-84. 2005.

Chakir, R. – “Spatial downscaling of agricultural land use data: an economic approach using cross entropy”. *Land Economics*, 85(2), pp. 238-251. 2009.

Fragoso, R.; Carvalho, M.L. – “Estimation of Farm Cost Allocation Coefficients at Farm Level Using an Entropy Approach”. *CEFAGE-UE Working-Papers*, 2011/21. 2011.

Fragoso, R.; Martins, M.B.; Lucas, M.R. – “Disaggregated soil allocation data using a Minimum Cross Entropy model”. *WSEAS Transactions on Environment and Development*, 9 (4), pp. 756-766. 2008.

Fraser, I. – “An application of maximum entropy estimation the demand for meat in the United Kingdom”. *Applied Economics*, 32, 45-59. 2000.

Garvey, E.; Britz, W. – “Estimation of Input Allocation from EU Farm Accounting Data using Generalized Maximum Entropy”. *CAPRI Working Paper* 02-01. U.Ireland & Bonn. 2002.

Golan, A.; Judge, G. – “Recovering and processing information in the case of underdetermined economic inverse models”. Unpublished Manuscript. University of California- Berkeley. 1992.

Golan, A.; Judge, G.; Robinson, S. - "Recovering information from incomplete or partial multisectorial economic data". *Review of Economics and Statistics*, 76:541-551. 1994

Golan, A. and J. Perloff (2002), 'Comparison of maximum entropy and higher-order entropy estimators'. *Journal of Econometrics* **107**(1-2): 195-211.2002

Golan, A.; Judge, G.; Miller D. – *Maximum Entropy Econometrics: Robust Estimation with Limited Data*. New York. John Wiley Editions. 1996a.

Golan, A.; Judge, G.; Perloff M. – "A Maximum Entropy Approach to Recovering Information From Multinomial Response Data". *Journal of the American Statistical Association*, Vol.91, n.434, 841-853. 1996b . Disponível em <http://www.jstor.org/stable/2291679>

Golan, A.; Perloff, M.; Shen, Z. – "Estimating a demand system with the non-negativity constraints: Mexican meat demand". *Review of Economics and Statistics*, LXXXIII:541-551. 2001.

Good, I.J. – "Maximum entropy for hypothesis formulation, especially for multidimensional contingency tables". *The Annals of Mathematics Statistics*, Vol. 34, N°. 3:911-387. 1963.

Hansen, H.; Surry, Y. – "Estimating the cost allocation for German agriculture: an application of the maximum entropy methodology". In: Conference paper, *46th Annual Conference of German Association of Agricultural Economists*. 2006.

Harris, R. – "Estimation of a Regionalized Mexican Social Accounting Matrix:using Entropy Techniques to Reconcile Disparate Data Sources". *TMD discussion paper n°97 WDC*. USA. 2002

Hazell, P.; Norton, R. - *Mathematical Programming for Economic Analysis in Agriculture*. NY, Mac Millan P.C. USA. 1986.

Howitt, R.; Reynaud, A. – "Spatial Disaggregation of Agricultural Production Data using Maximum Entropy". *European Review of Agriculture Economics*. Vol. 30, N°. 3, 359-387. 2003.

Howitt, R.; Msangi, S. – "Estimating Disaggregate Production Functions: An Application to Northern Mexico". Draft paper.2006.

Jaynes, E.T. – "Prior Information and Ambiguity in Inverse Problems". In: McLaughlin (ed.), *Inverse Problems, Providence RI: American Mathematical Society*. 151-166. 1984.

Jaynes, E.T. – "Prior Probabilities". *IEEE Transactions on systems science and cybernetics*. Vol. 4,n° 3, 227-241. 1968.

Jaynes, E.T. – “Information theory and statistical mechanics”. *Physics Review*.106, pp. 620-630. 1957a.

Jaynes, E.T. – “Information theory and statistical mechanics”. *Physics Review*. 108, pp. 171-190.1957b.

Jornal Oficial n.º L 109 de 23.6.1965, p. 1859, Regulamento 79/65/CEE de 15 de Junho de 1965

Jornal Oficial n.º L 148 de 5.6.1978, p. 1—34, Decisão 78/463/CEE: Decisão da Comissão, de 7 de Abril de 1978

Jornal Oficial n.º L 210 de 30/07/1981, p. 0001 – 0004- Regulamento (CEE) n.º 2143/81, de 27 de Julho de 1981

Jornal Oficial n.º L 335 de 13/12/2008, p. 0001 – 0024- Regulamento (CEE) n.º 1242/08, de 8 de Dezembro de 2008

Jornal Oficial n.º L 329 de 15/12/2009, p. 0001 – 0028- Regulamento (CEE) n.º 1200/09, de 30 de Novembro de 2009

Just, R.; Zilberman, D.; Hochman, E. – “Estimation of Multicrop production functions”. *American Journal of Agricultural Economics*. vol. 65, pp. 770-780. 1983.

Kitamura, Y; Stutzer, M. – “An information-theoretic alternative to generalized method of moment estimation”. *Econometrica*. vol. 66(4), pp. 861-874. 1997.

Lence, S.; Miller, D. – “Recovering output-specific inputs from aggregate input data: a generalized cross-entropy approach”. Agricultural and Applied Economics Association source *American Journal of Agricultural Economics*. vol. 80, nº4, pp. 852-867. 1998.

Léon, Y.; Peeters, L.; Quinqu, M.; Surry, Y. – “The use of maximum entropy to estimate input-output coefficients from regional farm accounting data”. *Journal of Agricultural Economics*.50: 425-439.1999.

Lips, M. – “Full product costs on base of farm accountancy data by means of maximum entropy”. Paper presented to the *International Association of Agricultural Economists Conference*. Beijing, China. August, 2009.

Love, H.A. "Conflicts Between Theory and Practice in Production Economics." *Amer. J. Agr.Econ.* 81,696-702. 1999.

Martins, M.B.; Fragoso, R.; Xavier, A. – “Spatial desagregation of Agricultural Data: A maximum Entropy Approach”. *JP Journal of Biostatistics*. Vol. 5,1, pp. 1-16. 2011.

Midmore, P. – “Estimating input-output coefficients from regional farm data: a comment”. *Journal of Agricultural Economics*. 41, pp. 108-111. 1990.

Moxey, A; Tiffin, R. – “Estimating linear production coefficients from farm business survey data: A note”. *Journal of Agricultural Economics*. 45, pp. 381-385.1994.

Offerman, F.; Berner, A. – “Estimation of production costs based on the German FADN”.In *FACEPA*. 2010.

Papalia, R. – “Generalized Maximum entropy Estimation of Spatial panel data interaction models”. *WSEAS Transactions on Business and Economics*. 14, Vol9, pp. 322-331.2010.

Paris, Q.; Howitt, R.E. – “An analysis of ill-posed production problems using maximum entropy”. *American Journal of Agricultural Economics*. n.80, pp. 124-138. 1998.

Peeters, L., Surry, Y. - Farm cost allocation based on the maximum entropy methodology- The case of Saskatchewan crops farms-Agriculture and Agri-Food Canada Strategic Policy Branch. Publication 2121/E, Project 02-011-tp, Ottawa,Ontario.2002.

Peeters, L.; Surry, Y. – “Maximum entropy estimation of a restricted Hildreth-Houck random-coefficients model”. In a Conference in Honor of Arnold Zellner: *Recent Developments in the Theory, Method, and Application of Information and Entropy Econometrics*: American University, Washington, D.C. 2003. Disponível em <http://www.american.edu/academic.depts/cas/econ>.

Peeters, L.; Surry, Y. – “Estimation d’un modèle à paramètres variables par la méthode d’entropie croisée généralisée et application à la répartition des couts de production en agriculture”. In : *Actes des Journées de Méthodologie Statistique*. 2005.

Pires, C.; Dionísio, A.; Coelho, L. – “GME versus OLS which is the best to estimate utility functions”. *CEFAGE-UE Working-Papers*. 2010/02. 2010.

Polomé, P.; Fernagut, B.; Harmignie, O.; Henry de Frahan, B. – “Multi-input Multi-output Farm-level Cost Function: A Comparison of Least Squares and Entropy estimators”. In: *11th Congress of EAAE proceedings*.Denmark. 2005.

Pukelsheim, J. – “The three Sigma Rule”. *The American Statisticians*. 48(4), pp. 88-91. 1994.

Rezek, J.; Campbell, R. – “Cost estimates for multiple pollutants: A maximum entropy approach”. *Energy Economics*. 29, pp. 503-519. 2006. Disponível em <http://www.sciencedirect.com>

Shannon, C. – “A Mathematical Theory of Communication_(1 and 2)”. *Bell Systems Tech.* 27, pp. 379-423 and 623-656. 1948.

Silva, R. – “Caracterização estatística de extremos sísmicos via distribuição generalizada de Pareto. Estudo de caso: João Camara-RN.”. Dissertação de mestrado-UFRGN-CCET-PPGMAE. 2008.

Soofi, E. S. – “Principal information theoretic approaches”. *Journal of the American Statistical Association*, Vol.95, pp. 1349-1353. 2000.

Van Delden, H.; Luja, P. “Integration of multi-scale dynamic spatial models for land use change analysis and assessment of land degradation and socio-economic processes”. In: *Conference on Soil Soil protection strategy - needs and approaches for policy support*, Poland. 2006.

support, Poland. 2006.

Xavier, A.; Martins, M.B.; Fragoso, R. – “Combined disaggregation of agricultural land uses, livestock numbers and crops’ production: an entropy approach”. *Advances in Mathematical and Computational methods*. pp 192-198. 2010.

Xavier, A.; Martins, M.B.; Fragoso, R. – “A minimum cross entropy model to generate disaggregated agricultural data at the local level”. In: *122nd EAAE Seminar-Evidence-Based Agricultural and Rural Policy Making:Methodological and Empirical challenges of Policy*. Ancona. 2011.

Zhang, X.; Fan, S. – “Estimation Crop-Especific Production Technologies in Chinese Agriculture: a Generalized Maximum Entropy Approach”. *American Journal of Agricultural Economics*. Vol.83, n.2, pp. 378-388. 2001. Disponível em <http://www.jstor.org/stable/1244680>

ANEXOS

Anexo 5.1.- Exemplo de uma parte de listagem do GAMS

GAMS Rev 237 WIN-VS8 23.7.2 x86/MS Windows 11/14/11 14:53:22 Page 1
General Algebraic Modeling System
Compilation

```

1  *Estimation of cost-allocation coefficients with the ME-tobit specificatio
n
2  *Sample of farms with technical economic orientation P1 arable crops
3  *from RICA Alentejo Region 2008 year
4
5  SETS
6  T farm index / 1 * 29 /
7
8  I costs index / 1 seeds and plants
9      2 fertilizer
10     3 chemicals
11     4 other costs with crops
12     5 margin/
13
14  k output index / 1 wheat
15     2 maize
16     3 riz
17     4 other cereals
18     5 legumes
19     6 horticultural crops
20     7 oilseeds
21     8 olive trees
22     9 permanent pastures
23    10 temporary pastures
24    11 forage maize
25    12 other forage/
26
27  M index parameter support /1*3/
28  N index error /1*3/
29 ;
30
31  ALIAS (i,j);
32  parameters
33  * Z(m) is the vector of support values for input/output coefficients
34
35  z(m)
36  /1  0
37   2  0.5
38   3  1
39  /
40  * VI(n) is the coefficients of error support values
41  VI(n) /1 -1
42   2  0
43   3  1/
44 ;
45  Table Y(t,k) output values in euros per crop activity

107 Table X(t,i)

139 table rcost(k,i) observed cost-allocation coeffi per output activity (MBS)

154 ;
155 PARAMETERS
156 X(t,i)
157 XYZ(t,i)
158 rcost(k,i)

```

```

159 XK(k,i)
160
161 *cost allocation coefficients
162 Alloct(k,i)
163 *estimated residuals
164 W(t,i,n)
165 sigma(i,j)
166 *predicted values of input costs
167 CostK(t,i) prediction of input costs in euros
168 CostKK(k,i)
169 RSquarr(i)
170 *normalized entropy
171 LnM
172 IK_LnM
173 IT_LnN
174 CRITP_ik(i,k)
175 CRITP
176 CRITU
177 PAD(k,i)
178 DIG
179 ;
180 *estimate of the standard error of the residuals
181 Parameters
182 Maxim(i)
183 Minim(i)
184 sdev(n,i)
185 V(n,i)
186 ;
187 maxim(i)=smax(t,X(t,i));
188 minim(i)=abs(smin(t,X(t,i)));
189 sdev(n,i)=(((maxim(i)-minim(i))**2)/12)**(1/2)*VI(n);
190 V(n,i)=3*sdev(n,i);
191 display sdev,V;
192
193
194 VARIABLES
195 OBJ;
196
197 POSITIVE VARIABLES
198
199 * probabilities vectors of cost allocation and of error terms
200 p(k,i,m)
201 alfa(k,i)
202 w1(t,i,n)
203 w2(t,i,n)
204 ;
205 *defining the variables that positive or equal to zero (tobit)
206
207 XYZ(t,i)=0$(X(t,i) LE 0)+X(t,i)$ (X(t,i) GT 0);
208 XK(k,i)=0$(rcost(k,i) LE 0)+rcost(k,i)$ (rcost(k,i) GT 0);
209
210
211 EQUATIONS
212 OBJECTIVE
213 ADD1(k,i)
214 ADD21(t,i)
215 ADD22(t,i,n)
216 ADD31(t,i)
217 ADD32(t,i,n)
218 CON1(t,i)
219 CON2(t,i)
220 SUMRESTER(K)
221 Valfa(K,i)

```



```

222
223 ;
224
225 *Objective function
226
227 OBJECTIVE.. OBJ =e= -sum(k,sum(i,sum(m,p(k,i,m)*log(1.e-4+p(k,i,m))))
228      -sum(i,sum(t$(XYZ(t,i) GT 0),sum(n,w1(t,i,n)*log(1.e-4+w1(t,i,n
229      ))))
230      -sum(i,sum(t$(XYZ(t,i) LE 0),sum(n,w2(t,i,n)*log(1.e-4+w2(t,i,n
231      ))));
232
233 *properties of probabilities
234
235 ADD1(k,i).. sum(m,p(k,i,m)) =e= 1;
236
237 ADD21(t,i)$ (XYZ(t,i) GT 0).. sum(n,w1(t,i,n)) =e= 1;
238 ADD22(t,i,n)$ (XYZ(t,i) LE 0).. w1(t,i,n) =e= 0;
239
240 ADD31(t,i)$ (XYZ(t,i) LE 0).. sum(n,w2(t,i,n)) =e= 1;
241 ADD32(t,i,n)$ (XYZ(t,i) GT 0).. w2(t,i,n) =e= 0;
242
243 CON1(t,i)$ (XYZ(t,i) GT 0).. XYZ(t,i) =e= sum(k,sum(m,p(k,i,m)*z(m))*y(t,k)
244      )
245      +sum(n,w1(t,i,n)*v(n,i));
246 CON2(t,i)$ (XYZ(t,i) LE 0).. sum(k,sum(m,p(k,i,m)*z(m))*y(t,k))
247      +sum(n,w2(t,i,n)*v(n,i)) =l= 0;
248
249 SUMRESTER(K).. sum(i,sum(m,p(k,i,m)*z(m))) =E= 1;
250
251 valfa(K,i).. sum(m,p(k,i,m)*z(m)) =E= alfa(k,i);
252
253

```


Anexo 5.2.- Análise descritiva dos desvios absolutos médios

Tabela A 5.1.- Análise descritiva do DAM em MEG nas Culturas Arvenses

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	25,1	55,3	2,2	35,5	49,6
Máximo	292,0	670,0	615,0	413,3	77,4
Mínimo	0,0	0,0	0,0	0,0	17,2
Média	67,6	129,7	103,0	96,9	51,6
Desvio padrão	95,4	207,4	246,2	141,5	58,2
Coefic. variação	141,2	159,9	239,1	146,0	112,8

Fonte: Resultados do Modelo

Tabela A 5.2.- Análise descritiva do DAM em ECG nas Culturas Arvenses

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	28,1	22,2	47,6	18,8	6,0
Máximo	56,8	137,4	203,0	165,0	45,5
Mínimo	0,0	0,0	0,0	0,0	0,6
Média	24,8	37,2	62,7	34,1	10,9
Desvio padrão	18,6	40,1	72,6	47,3	12,4
Coefic. variação	74,9	107,7	115,8	138,5	114,6

Fonte: Resultados do Modelo

Tabela A 5.3.- Análise descritiva do DAM em MEG nas Culturas Permanentes

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	0,0	32,5	0,0	90,6	97,1
Máximo	74,6	77,7	1370,1	99,6	478,4
Mínimo	0,0	1,1	0,0	0,0	28,0
Média	26,6	34,4	277,9	68,6	163,3
Desvio padrão	36,9	29,3	610,6	41,4	182,3
Coefic. variação	138,5	85,3	219,7	60,4	111,6

Fonte: Resultados do Modelo

Tabela A 5.4.- Análise descritiva do DAM em ECG nas Culturas Permanentes

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	0,0	73,0	0,0	2,0	6,0
Máximo	74,6	152,3	1117,2	4,0	40,9
Mínimo	0,0	27,3	0,0	0,0	0,1
Média	18,6	76,9	255,7	2,0	12,6
Desvio padrão	32,3	47,8	486,6	1,6	16,4
Coefic. variação	173,6	62,1	190,3	79,1	130,3

Fonte: Resultados do Modelo

Tabela A 5.5.- Análise descritiva de DAM em MEG nos Herbívoros componente Vegetal

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	50,3	311,1	294,4	266,7	35,7
Máximo	161,9	2563,6	843,8	1600,0	78,9
Mínimo	37,3	24,2	32,6	94,1	2,2
Média	75,0	633,2	367,3	490,2	35,0
Desvio padrão	58,7	880,9	343,2	628,6	27,1
Coefic. variação	78,3	138,1	93,4	128,3	77,2

Fonte: Resultados do Modelo

Tabela A 5.6.- Análise descritiva de DAM em ECG nos Herbívoros componente Vegetal

	Sem./Planta	Fertilizantes	Pesticidas	Outros Custos	Margem Bruta
Mediana	11,0	28,7	6,5	44,5	19,4
Máximo	79,4	86,3	87,3	99,9	82,4
Mínimo	0,0	0,0	0,0	0,0	0,0
Média	25,7	34,3	21,7	66,9	24,8
Desvio padrão	31,6	31,1	31,9	35,2	27,0
Coefic. variação	122,8	90,7	147,3	53,4	109,1

Fonte: Resultados do Modelo

Tabela A 5.7.- Análise descritiva de DAM em MEG nos Herbívoros componente Animal

	Concentrados	Forragens	Outros custos	Margem Bruta
Mediana	74,8	287,0	116,4	49,3
Máximo	152,1	460,0	264,3	66,0
Mínimo	0,0	68,8	22,3	49,0
Média	75,4	275,7	129,8	53,4
Desvio Padrão	87,1	161,3	106,3	8,4
Coefic. variação	115,5	58,5	81,8	15,8

Fonte: Resultados do Modelo

Tabela A 5.8.- Análise descritiva de DAM em ECG nos Herbívoros componente Animal

	Concentrados	Forragens	Outros custos	Margem Bruta
Mediana	25,8	35,7	210,7	38,6
Máximo	51,9	73,0	427,9	57,4
Mínimo	0,0	11,4	66,9	37,0
Média	25,9	38,9	229,1	42,9
Desvio Padrão	29,9	25,7	156,7	9,8
Coefic. variação	115,5	66,1	68,4	22,8

Fonte: Resultados do Modelo